THE UNIVERSITY *of* EDINBURGH

CHARACTERISATION OF DISSOLVED ORGANIC MATTER IN WATER TREATMENT USING ULTRA-HIGH RESOLUTION TECHNIQUES

THE UNIVERSITY *of* EDINBURGH

ALAN JAMES SMITH

Doctor of Philosophy
EaSTCHEM School of Chemistry
The University of Edinburgh

September 2020

DECLARATION

I declare that this thesis has been composed solely by myself and that the work herein has been conducted by myself or when part of a research group, my contribution has been clearly indicated. The work presented has not been submitted for any degree or professional qualification other than this.

Signed: Alan James Smith

Date: 30/09/2020

LAY SUMMARY

Dissolved organic matter (DOM) is a highly complex mixture of organic compounds resulting from the degradation of plant, animal and microbial life. DOM is present in most if not all of the water sources on earth. This ubiquitous mixture affects many aspects of our lives, including the production of drinking water.

DOM has the potential to negatively impact every stage of water treatment. Whether it be through increasing the volume of chemicals needed to treat the water or reacting directly with chemical disinfectants, producing by-products that are potentially harmful to consumers.

This project uses ultra-high resolution analytical techniques, namely, nuclear magnetic resonance (NMR) and Fourier transform ion-cyclotron resonance mass spectrometry (FT-ICR-MS) to investigate how the composition of DOM changes throughout various stages of water treatment in plants producing potable water and in a laboratory setting.

Water collected at different stages of water treatment in small-scale treatment works was analysed to evaluate modern treatment techniques such as suspended ion exchange (SIX) and granulated activated carbon (GAC) filtration. We show that two forms of ion exchange (IEX) treatment were selective in the removal of compounds that are known to result in the production of regulated disinfection by-products. The IEX treatments were also shown to be more effective than the GAC filtration, resulting in a higher quality water source.

In-lab experiments were performed to investigate the photocatalytic degradation of DOM over time, characterising the degraded products as well as those produced.

Initial experiments showed that chloramination of DOM using radicals to mineralise the organic compounds. These radicals are produced by irradiating catalysts by ultraviolet/visible light. Different catalysts and irradiation wavelengths were investigated. FT-ICR-MS and NMR analyses characterised the degraded DOM as well as molecules produced during the photodegradation process. We were able to show that a modified titanium oxide catalyst is more efficient in this process (compared to unmodified) and can harvest relatively lower energy light from the edge of the visible spectrum. This is important, as these wavelengths are part of solar irradiation, its use would eliminate the need for the high energy UV lamps that are currently used.

The final part of this project was a study of the by-products formed during chloramination of DOM. Chloramination is a method of disinfection, the chloramines used in this process have the potential to react with DOM. Our initial experiments showed that the chloramination of DOM resulted in a mixture of by-products so complex that it would be impractical to study them. Small model molecules were therefore used; however, these experiments still produced hundreds of disinfection by-products. A molecule incorporating a fluorine nucleus, was therefore used as a model compound. Fluorine is a very convenient tag that is easily followed by NMR. To enable this research a suite of NMR experiments was designed to extract necessary structural information from a complex mixture of products, without the need to separate individual molecules. Structures of a dozen by-products were determined, laying the foundation for the future of NMR investigations into the reaction between DOM and disinfectants.

ABSTRACT

Dissolved Organic Matter (DOM) is an extremely heterogeneous complex mixture, consisting of thousands of chemical species. The composition of DOM greatly influences the potable water production process. Current methods employed by industry for the characterisation of DOM are relatively low resolution; namely Total Organic Carbon (TOC) analysis or molecular weight analysis via Liquid Chromatography – Organic Carbon Detection (LC-OCD). In this thesis, high-resolution spectrometric and spectroscopic techniques were used to investigate the complexity of DOM and to characterise changes that occurred throughout various treatment processes.

A pilot plant evaluating the use of a Suspended Ion Exchange (SIX) system, followed by coagulation and ceramic membrane filtration, for the removal of DOM was sampled and investigated. Nuclear Magnetic Resonance (NMR) and negative mode electrospray – Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (ESI (-) FT-ICR-MS) were used to characterise the organic species present within the inlet water, post ion-exchange and the final outlet water. The samples from this pilot plant were also compared to those of the existing, full-scale water treatment works present on the same site. This study found via mass spectrometry that the SIX treatment was non-selective in the removal of compounds, with the composition of samples being highly similar to those of the raw water. The coagulation and ceramic membrane filtration treatment, however, was selectively removing aromatic and phenolic species. These findings were also corroborated with the use of NMR and Fourier Transform Infrared Spectroscopy (FT-IR). Compounds with these characteristics have been shown to result in the production of by-products in later treatment stages, so their removal is desirable.

A second pilot plant was investigated to evaluate the use of granular activated carbon (GAC) and ion-exchange (IEX) filtration as post coagulation treatments for the removal of DOM over the course of 6 months. Throughout the 6-month period, the IEX consistently outperformed the GAC treatment in terms of number of species removed, reduction of aromatic compounds and performance stability over time. The GAC treatment however, also resulted in the removal of aromatic compounds and is a less expensive system to implement and maintain. DOM has been shown to negatively impact many of the processes used to create potable water, increasing coagulation loads, reducing the lifetime of any

filtration systems, and creating by-products in the disinfection stage. This makes DOM removal essential for improving the quality of the produced drinking water and the efficiency of a treatment works.

Photocatalysis is a highly promising method for the oxidation of DOM and its mineralisation. TiO_2 is a well-known photocatalyst capable of degrading DOM, when activated by UV light. The composition of Suwannee River Fulvic Acid (SRFA), a common international DOM standard, throughout photocatalytic degradation under various wavelengths was characterised across a period of three hours and three irradiation wavelengths (370 nm, 410 nm and white LED). The performance of pristine TiO_2 was compared to that of a TiO_2-based catalyst that had been doped with bismuth. We show that the mechanism of degradation is highly likely to be the same for both catalysts and that the performance of the doped catalyst is superior to that of pristine TiO_2; doping of TiO_2 has enabled more efficient utilisation of softer irradiation. Based on a detailed analysis of ESI (-) FT-ICR-MS data, the effects of photocatalysis were monitored in terms of molecular weight distribution, double-bond equivalent, aromaticity, oxygen numbers and a susceptibility of compound classes. Chemometric analysis of ^{1}H NMR data highlighted the existence of long chain fatty acids as products of the photolytic degradation. This work represents the most detailed molecular level analysis of photocatalytic degradation of DOM to date.

Disinfection of potable water is essential to providing the public with a safe drinking source. DOM has the potential to react with chemical disinfectants such as chlorine, the most widely used primary disinfectant. This results in the production of a variety of disinfection by-products (DBPs), some of which are regulated. To reduce the production of DBPs, water suppliers are exploring alternative disinfection processes. One process gaining traction is chloramination. In this thesis, we explore chloramination via high-resolution NMR, employing ^{19}F as an NMR probe into the complexity of DBPs. To effectively achieve this, existing ^{19}F NMR experiments were optimised, and new pulse sequences developed. These include methods for establishing the ^{1}H - ^{19}F and ^{19}F - ^{13}C correlation utilising far-reaching proton-fluorine and carbon-fluorine scalar coupling constants. The obtained coupling constant and correlated ^{1}H, ^{13}C and ^{19}F chemical shifts were used to propose structures of a dozen chloramination DBPs produced from 3-Fluoro-4-hydroxybenzoic acid, a compound with functional groups known to be prevalent in DOM. Decarboxylation, dehydroxylation, chlorination and an addition of nitrogen were observed. The developed methodology will assist in developing chloramination as an industrial process.

PUBLICATIONS

The following is a list of publications and potential publications resulting from the work conducted during this project. If the work is included in this thesis, this is declared at the start of each chapter.

Published

(1) Smith, Alan J.R and Moore, Graeme and Semiao, Andrea and Uhrín, Dušan. *Environmental Science: Water Research & Technology*, 2020, **6**, 1495-1504.

(2) Bell, Nicholle G.A and Smith, Alan J.R *et al. Nature, Scientific Reports* 2020, **10**, 1484.

Unpublished

(1) Photocatalytic Degradation of Dissolved Organic Matter Tracked by FT-ICR-MS and NMR Spectroscopy

(2) NMR Experiments for the Analysis of Complex Mixtures of Mono-Fluorinated Compounds. Investigation of Chloramination Disinfection By-Products

ACKNOWLEDGMENTS

First and foremost, I need to thank my supervisor Dušan Uhrín. You have given me guidance and instruction when desired and patience when sorely needed. Having gone through several problems throughout this project, you have helped me from start to finish.

I would also like to thank Andrea Semiao for her help, support and for permitting me the use of her membrane lab, without which I would probably still be concentrating my water samples.

To Scottish Water for funding the project, thanks to Paul Weir, Graeme Moore and Scott Lambie for making me feel welcome and allowing me to sample the sites presented in this thesis.

My sincere thanks go out to the staff at the University of Edinburgh namely, Juraj Bella, Lorna Murray, Faye Cruickshank and Logan Mackay. I have learned much from you all regarding use and maintenance of the instrumentation, through many a testing situation, may it remain useful throughout my career.

Undertaking a PhD is no easy feat, it is one which presents many challenges, some of which I may not have gotten over if not for the good friends I have made throughout the duration of this project. To John, Natalia, Nicholle, Elaine, Gian, Claire, Marc and Richard, thank you all for making the office not only a great place to work, but somewhere I have enjoyed spending my time, I wish you all the very best.

To Will, Lorna and Ariana, you made me feel very welcome from day one, were always there to answer the regular call of the pub. You introduced me to many things that shall remain with me, a love of whisky, wonderful beer and great cheese (still, no blue thanks). You inspired the (amateur) chef in me with trips to Fazenda, Chophouse and Newhaven Fish Market. I'm glad to call you all my friends, without you my experience here would have been more tame and far less enjoyable.

To my Mum and Dad, your continued support and advice throughout this time has been invaluable. Visiting me so often has really meant a lot, experiencing the best that Edinburgh has to offer with me, love you both.

Saving the best for last. To Lucy, these past few years have not been easy, especially the constant travelling (over 15,000 miles by train!). You have been there with me through the good times, gotten me through some bad, I couldn't

have done this without your continued love and support, I will be forever grateful for your patience.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

AOP	Advanced oxidation process
AI	Aromaticity index
API	Atmospheric pressure ionisation
APPI	Atmospheric pressure photoionisation
BIRD	Bilinear rotation decoupling
CRAM	Carboxyl rich alicyclic molecules
CSA	Chemical shift anisotropy
CSSF	Chemical shift selective filter
CID	Collision induced dissociation
COSY	Correlation spectroscopy
DIPSI	Decoupling in the presence of scalar interactions
DOSY	Diffusion ordered spectroscopy
DBPs	Disinfection by-products
DOM	Dissolved organic matter
DBE	Double bond equivalence
ESI	Electrospray ionisation
FT	Fourier transform
FTICR	Fourier transform Ion Cyclotron Resonance
FID	Free induction decay
FWHM	Full width half maximum
FA	Fulvic acid
GAC	Granular activated carbon
HAAs	Halo acetic acids
Hz	Hertz
HSQC	Heteronuclear Single Quantum Coherence
HMBC	Heteronuclear Multiple Bond Correlation
HPSEC	High pressure size-exclusion chromatography
HA	Humic acid

INEPT	Insensitive nuclei enhanced polarisation transfer
IEX	Ion-exchange
KMD	Kendrick mass defect
LDI	Laser desorption ionisation
MS	Mass Spectrometry
MALDI	Matrix assisted laser desorption ionisation
MF	Microfiltration
NF	Nanofiltration
NOM	Natural organic matter
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser effect
POM	Particulate organic matter
ppm	Parts per million
RF	Radiofrequency
RO	Reverse osmosis
S/N	Signal-to-noise
SOM	Soil organic matter
TOF	Time of flight
TOCSY	Total correlation spectroscopy
TOC	Total organic carbon
THMs	Trihalomethanes
UF	Ultrafiltration

Part I

INTRODUCTION

INTRODUCTION

1.1 WHAT IS NATURAL ORGANIC MATTER (NOM)?

Natural Organic Matter (NOM) is an incredibly rich, complex mixture of organic material that originates from a multitude of sources, such as plant and animal decomposition, microbial activity and chemical degradation. As NOM components stem from living organisms, its composition not only depends on the local fauna and flora but also the environmental conditions. This causes NOM to consist of an extremely large number and variety of chemical constituents, making it inherently challenging to study.

Terrestrial organic matter, also known as Soil Organic Matter (SOM), is formed through a collection of processes known as humification, primarily from plant material. Biotic decay of plant material occurs mainly via microbial degradation, those molecular species which result in a higher energetic yield or provide greater nutrition are utilised first (carbohydrates and/or nitrogen containing species), resulting in the concentration of more recalcitrant material overtime. As microorganisms utilise these plant chemical constituents, they are also producing their own metabolites, which contribute to the chemical diversity of the SOM. Abiotic processes are also likely to occur as these chemical constituents react with one another and oxygen.

Due to the complexity of NOM, it is often broken down into smaller operationally defined subcategories:

i) Fulvic Acids (FA) - Fraction which is soluble at all pH values

ii) Humic Acids (HA) - Only soluble at high pH values

iii) Humin - Insoluble at all pH values

The NOM found within aquatic systems is often a product of soil runoff, with rain and groundwater carrying both dissolved NOM and particulates to water systems. Although there are reported cases of aquatic NOM being found in water systems that have no terrestrial source of water, an example being Pony Lake, a coastal pond on Ross Island, Antarctica, where the NOM present is thought to be exclusively of microbial origin.[1] Once dissolved, NOM can be further categorised as either Dissolved Organic Matter (DOM) or Particulate

Organic Matter (POM), the former being NOM that can pass through a filter pore size of 0.45 μm, the latter being NOM which cannot pass through this filter. All of these operationally defined categories represent highly heterogeneous mixtures consisting of molecules that cover many chemical classes.

As mentioned earlier, NOM is composed of a mixture of known biomolecules and partially decomposed/degraded material; these are often termed non-humic substances and humic substances (HS), respectively. The chemical structures associated with humic substances has been a difficult and controversial area of study. Early research based upon filtration or centrifugation through differing pore size membranes suggesting NOM was composed of structures across a vast mass range of 0.1-100 kDa,[2,3] with an average molecular weight of up to 50 kDa leading many to believe that NOM was composed of polymeric constituents. One theory stated that these polymers were formed via the Maillard reaction, a reaction between carbohydrates and amino acids shown to produce large polymeric species.[4]

This theory began to unravel with the advancement and application of chromatographic and spectroscopic techniques. Using High Pressure Size-Exclusion Chromatography (HPSEC) researchers showed that the apparent molecular weight distribution of NOM could be lowered by decreasing the sample pH, and increased when the pH was raised.[5] The theory behind this was that the molecules were aggregating in solution via hydrophobic forces and hydrogen bonding.[6] Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FT-ICR-MS) studies also showed that NOM samples possessed molecules between 100-1000 Da, much lower than was originally predicted. This area however remains of interest, for example a study published in 2019 highlighted, using high-resolution MS and HPSEC, that there exists a significant fraction of high molecular weight, UV active NOM which is not observable via their MS system.[7]

The structures of humic substances are an important field of study for many reasons, they are ubiquitous to all soils, sediments and water bodies and have the potential to influence many geochemical processes. They are an important component of the global carbon cycle, contributing to the fixation of carbon, acting as a carbon sink. The chemical moieties present on HS have been shown to control the binding of metals, such as arsenic, uranium and aluminium via complexation.[8–10] Having a better understanding of the structural components present within HS, will enable better assessment of risks associated with heavy metals and other contaminants known to bind to humic substances.

As discussed further in this thesis, NOM plays an important role throughout the production of drinking water, it is not a controlled parameter of potable

water, but it does heavily influence aspects which are regulated. To that end, this thesis aims to utilise ultra-high resolution analytical techniques such as nuclear magnetic resonance (NMR) and Fourier transform ion-cyclotron resonance mass spectrometry (FT-ICR-MS) to provide a greater insight into the composition of NOM, and how the composition is altered throughout various treatment processes.

1.2 WATER TREATMENT DURING POTABLE WATER PRODUCTION

1.2.1 *Overview*

Access to a reliable and clean water supply is one of the most influential aspects in the health of a population, improved water supply and sanitation can boost economic growth and contribute greatly to poverty reduction. Contaminated water is linked to the transmission of many diseases via contact and ingestion, such as cholera and typhoid.

It is not only pathogens that need to be remediated or removed from the drinking supply in order to produce clean, potable water. As freshwater travels through the environment to reservoirs or percolates through the soil, it comes into contact with a variety of environments and geologies, picking up metals, minerals and organics of natural and anthropogenic origin along the way. Many of these constituents can be harmful to not only the consumer, but also the distribution network or containment facilities. The overarching goal of potable water treatment is to prevent or reverse the contamination of water in order to provide a population with a biologically and chemically safe water source. The drinking water quality regulator (DWQR) for Scotland currently tests and controls 51 different substances/parameters; some of their *key* parameters are shown in Table 1.1.[11]

In the subsections that follow, a short description of the major/most common treatment techniques used to produce potable water will be given. These descriptions will be mainly focused on how the treatment processes work and their influence on the fate of the organic matter present, as that is the direct focus of this thesis.

Table 1.1: Table showing various parameters controlled by the Drinking Water Quality Regulator for Scotland.[11]

Parameter	Concentration/Value
Aluminium	200 $\mu g/L$
Coliforms	0/100 ml
Copper	2 mg/ml
E.Coli	0/100 ml
Iron	200 $\mu g/L$
Lead	10 $\mu g/L$
Total Trihalomethanes	100 $\mu g/L$
Turbidity	4 NTU at tap

1.2.2 *Coagulation and Flocculation*

Coagulation was one of the first processes used to clean drinking water and is still a cornerstone for many treatment works today. The entire purpose of coagulation is to destabilise small particles which are suspended or remove dissolved species from solution. These destabilised particles are then encouraged to collide in a process known as flocculation, the collisions cause these particles to agglomerate and grow in size, becoming flocs. When the particles to be removed are larger, they are easier and less expensive to filter out, hence why coagulation remains extremely popular.

The destabilisation of particles is caused by chemical additives known as coagulants, the most common of these are inorganic metal salts, with aluminium sulfate being the most common due to being inexpensive and widely available (most common coagulants are shown in Table 1.2). When these salts are added to water they react via hydrolyses reactions, forming metal hydroxides; these charged products and the NOM are removed via a combination of several aggregate forming mechanisms such as charge neutralisation, adsorption and complexation.[12]

Table 1.2: Table of the most commonly used coagulants and their molecular formulae

Coagulant	Formula
Aluminium sulfate	$Al_2(SO_4)_3.14H_2O$
Polyaluminium chloride	$Al_2(OH)_5Cl$
Iron chloride	$FeCl_3$
Iron sulfate	$Fe_2(SO_4)_3$

In water conditions where the metal hydroxide precipitation is minimal, the cationic metal species interact with the anionic NOM components via electrostatic forces to form insoluble charge neutral products. This has been shown to be the dominant mechanism for NOM removal.[13] In waters where the pH causes the rapid formation of highly insoluble hydrolysis products, the major mechanism shifts to adsorption, where NOM components adsorb onto the surface of these products as they precipitate.[14] The third mechanism, complexation, occurs when the metal cation forms a soluble complex with the NOM, it only drops out of solution when the metal-NOM complex reaches saturation and precipitates.

The mechanism of removal has also been shown to differ depending on the composition/fraction of the NOM being removed. The main mechanism for high molecular weight NOM removal is expected to be charge neutralisation, whereas for low molecular weight, this is expected to be adsorption onto the coagulant metal surface. It is generally easier to remove higher molecular weight species and also those which are relatively hydrophobic. A problematic fraction are species which are hydrophilic and relatively low molecular weight.[13,15]

Under certain water conditions there exists a dominant removal mechanism as highlighted above, however in reality most/all of these mechanisms will contribute to some extent to NOM removal.[12,14]

After coagulation, particles are now destabilised and are termed micro-flocs, these can range in size (1 - 10 μm), structure and strength depending on the mechanism of removal.[16] In order to effectively remove these flocs, they either need to be filtered out (expensive) or allowed to fall to the bottom of the tank producing a sediment (cheaper). In order for the micro-flocs to drop to the bottom of the tank they need to grow in size; this occurs when the particles collide with and adhere to one another in a process known as flocculation. This can happen via Brownian motion of the water (perikinetic flocculation)

however, this is a slow process and is usually sped up through forced fluid motion (orthokinetic flocculation), commonly performed with mechanical stirring.

The effectiveness of coagulation is also dependant on factors such as pH, temperature, coagulation type and dosage of coagulant. There has been significant research into the coagulation properties of ferric-based coagulants, mainly due to the health risks associated with aluminium. Ferric-based coagulants have been shown to remove NOM more effectively than aluminium based coagulants and perform better at lower temperatures.[17–20] Ferric coagulants may perform better due to their tendency to produce floc of a greater size, which increases the chance of collisions and therefore NOM removal.[21] The effectiveness of a coagulant has been linked to the distribution of the resulting charged hydrolysis species. Those with higher overall surface charge can interact with more NOM prior to becoming neutralised, for this reason there are several studies entertaining the possibility of alternative metals, such as zirconium and titanium.[21–23]

In practice, the optimum coagulant dose and pH is usually determined through a series of jar-tests wherein samples of water are dosed with a variety of coagulant dosages under rapid mixing conditions. The water is then left for a standard period of time under a slower stirring speed allowing the formation of flocs. This is followed by a time of no mixing to allow for sedimentation before the turbidity of the water is measured to determine the optimum conditions. This however is a time costly process with multiple experiments needing to be performed and is not ideal when a treatment works has to deal with a fluctuating water source. To combat this many water works now utilise zeta potential, mostly in a lab setting, however there are online systems available. The zeta potential is a parameter that indicates the charge of a particle and can be used to inform coagulation dose.[24,25] To measure the zeta potential, a sample of water within a cell is subjected to an applied electric field. As the particles within the cell move towards the electrodes they cause scattering of a laser beam. The frequency of the scattering is related to electrophoretic mobility (proportional to particle speed). This frequency is determined at multiple voltages to calculate the zeta potential, i.e. a measure of the electrostatic repulsion or attraction of the system/sample. Systems with high zeta potentials, such as > 60 mV are reported to have excellent stability, whereas, those with a potential closer to 0 mV will be extremely unstable and are more likely to rapidly coagulate when a coagulant is added. The optimal coagulant dose can be inferred from zeta potential vs coagulant concentration plots.

1.2.3 *Filtration and Adsorption*

There are two main types of filtration used to produce potable water, media and membrane filtration. Filtration through media, such as sand, is relatively cheap and one of the most widely used water treatment techniques, with the first municipal sand filter being introduced in 1804, being installed in Paisley. In comparison, membrane filtration is relatively new, being introduced to the UK around 1994, it is generally more expensive than media filtration but there are very few pollutants that cannot be removed, given the appropriately sized membrane.

The efficacy of media filtration is largely dependant on the media used. Relatively inert materials such as sand are the most common because they are readily available, cheap and easy to clean. Sand filtration can only remove particulate matter; therefore it is commonly used in combination with coagulation, straining out the flocs as the water flows through. This does mean however, that the sand needs to be cleaned when it becomes fouled; this is usually done with backwashing or pressure washing.

More recently, there has been significant investment in alternative media which target organic species; they are capable of interacting with soluble components within the water via adsorption mechanisms. One such medium is Granular Activated Carbon (GAC), a carbonaceous material with a huge internal surface (500 - 2000 $m^2\ g^{-1}$), usually produced by pyrolysis (600 - 900 $^\circ$C) of a carbon based material in an inert atmosphere. Like all adsorbents, GAC has a limited adsorption capacity and once reached, the media needs to be regenerated (reheated) or replaced. Metals present within the water can also accumulate within the pores of GAC, reducing adsorption sites for NOM. Moreover these metals cannot be removed by reheating the adsorbent. In order to restore the GAC to its original pore characteristics it must be acid washed.[26]

Different fractions of NOM compete for adsorption sites; it has been reported that lower molecular weight fractions adsorb to a greater extent. This may be due to having increased access to the internal structure, i.e. more surface area.[27,28] The adsorption of NOM to the GAC has been shown to be heavily influenced by pH, with a greater NOM removal being observed in more acidic conditions. It is thought that at a low pH an abundance of protons become adsorbed to the surface of the GAC, this in turn attracts the deprotonated moieties of the NOM, such as carboxyl and phenolic groups.[29]

Another effective method from the media filtration category, for removing NOM from drinking water is ion exchange (IEX). As a large fraction of NOM can be characterised as having acidic sites, they will exist in equilibrium

between a protonated (neutral) and non-protonated state (negatively charged). Ion-exchange processes take advantage of this by using a basic resin which exchanges a counter ion (usually chloride) with the charged NOM molecule, it should be noted that acidic resins are also available just less amenable to the removal of NOM. As IEX is also an adsorption process, the resin will need regenerating, however this is often more economical than with GAC and involves flushing or soaking the resin in a saturated solution of the counter ion (NaCl). IEX and coagulation have been shown to have complementary NOM removal characteristics so are commonly paired together; the former being shown to prefer lower molecular species, the latter, high molecular weight species.[30]

A popular, commercially available IEX resin known as MIEX® (Orica Australia Pty, Ltd), is a resin with magnetic properties that promotes aggregation of the resin particles, enhancing resin recovery in the system. Studies have shown high NOM removal (85-90%) when MIEX is held in contact with a NOM solution for 10 to 20 minutes,[31] and up to 87% removal of dissolved organic carbon (DOC) when implemented prior to coagulation.[32] Implementing IEX to remove NOM before coagulation reduces the amount of coagulant needed, reducing the input of metals such as aluminium and iron.[33] Alternatively, IEX could be used post coagulation to maximise the resins potential in removing only species the coagulation could not.

There are however, some disadvantages to using an IEX system. When high levels of phosphates are present in the raw water, they will be exchanged onto the resin as well. This can cause the growth of a biofilm which leads to blinding of the resin and ultimately the need for replacement.[34] As mentioned previously, IEX systems are regenerated using high concentrations of counter ions, usually brine. After regeneration this waste brine solution which also possesses high concentrations of NOM either needs to be disposed of or recycled, increasing running costs.[35]

Membrane filtration processes are increasingly being utilised for the removal of NOM, these techniques rely on a thin layer of a semipermeable material that is capable of removing substances based on either their physical or chemical properties. There exists four main size classes of membranes applicable to the production of potable water; this coupled with a wide range of membrane materials allows for the tailoring of a membrane facility to the water requiring treatment. The four size classes are microfiltration (MF), utrafiltration (UF), nanofiltration (NF) and reverse osmosis (RO). Their respective pore sizes are given in Table 1.3.

Table 1.3: Filtration membrane categories and their effective pore sizes.

Membrane	Pore Size (μm)
Microfiltration	0.1
Ultrafiltration	0.01
Nanofiltration	0.001
Reverse Osmosis	0.0001

MF and UF membranes with their relatively large pore sizes are termed porous membranes, meaning they operate like sieves, where the separation is purely mechanical via size exclusion. These membranes are more amenable to the reduction of turbidity through the removal of larger suspended particles and pathogens (*E. coli* width is 0.25 - 1 μm) but are capable of removing NOM when used in conjunction with a coagulant. The NF and RO membranes are produced from a very dense material that operates via a combination of mechanical and physicochemical relationships, such as charge repulsion, with the analyte in question. This provides these membranes with a much smaller *effective* pore size, with NF membranes being capable of removing fractions of dissolved organic matter and RO membranes being able to remove most organic material and dissolved salts.

With all membranes, the water passes through the membrane under an applied pressure either via dead-end or cross-flow processes. This results in the production of a clean stream of water, which can proceed on to further treatment, but also the production of a concentrated solution on the feed side of the membrane. This concentrate must be disposed of, increasing costs. It can also cause issues with the membrane such as blinding or biofouling, both requiring either replacement of the membrane, the use of a cleaning solution or mechanical scouring.

There have been many studies on the efficiency of NOM removal through membrane filtration. One study compared two types of UF membranes with three types of NF membranes on their NOM removal capabilities across ten different water sources. They found that on average the UF membranes removed 47-50% of the measured dissolved organic carbon (DOC), whereas the NF membranes, on average, were able to remove 86-90%. The permeate water was then subjected to trihalomethane formation potential (THMFP) tests, and showed that the water treated with NF membranes resulted in the formation

of significantly lower concentrations of disinfection by-products (discussed in next section).[36]

Reverse osmosis membranes have been shown to be capable of removing almost all organic matter from the water, with one group reporting DOC retention of $> 90\%$.[37] It is this high retention that leads many researchers, including myself, to utilise RO for the isolation of NOM for further studies. There exists a very common NOM standard created by the International Humic Substances Society (IHSS) which is produced using RO filtration. When creating a NOM concentrate, the high retention of RO also has some drawbacks such as the concentration of inorganic species. These are detrimental to many analytical techniques, such as mass spectrometry for example. In order to overcome this, RO is often coupled with electrodialysis, a technique that uses electrodes to draw ions through anionic/cationic membranes, resulting in the reduction of ions from one solution (diluate) and the concentration of ions in another (concentrate).

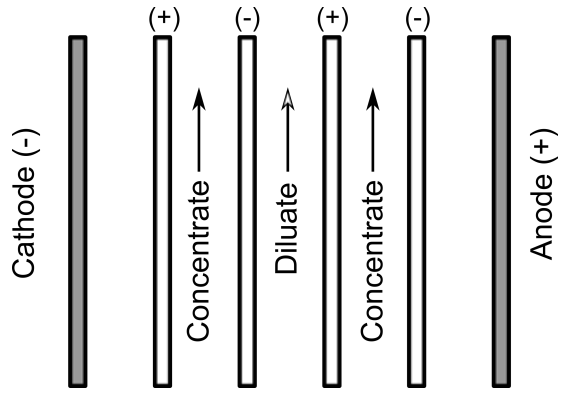


Figure 1.1: Schematic of an electrodialysis system with alternating cationic/anionic membranes.

This combined approach has been reported to result in the concentration of organic species in NOM, with only an average loss of 3% total organic carbon.[38,39]

1.2.4 *Disinfection*

Potable water must be disinfected in order to eliminate pathogenic organisms and ensure biological safety to the consumer. Disinfection is usually done by chemical means through the use of disinfectants, the most common of which is chlorine, which can be supplied either as a liquefied gas (Cl_2) or as a salt (NaOCl, $Ca(OCl)_2$).

Chlorine dissolves in water to form hydrochloric acid, hypochlorous acid and the hypochlorite ion, as per Equation 1.1.

$$\begin{aligned} Cl_2 + H_2O &\rightarrow HCl + HOCl \\ HOCl &\rightarrow H^+ + OCl^- \end{aligned} \tag{1.1}$$

At pH values common to surface waters, the predominant species is hypochlorous acid; the more active moiety, present above pH 9, the hypochlorite ion is the dominant species.[40] Hypochlorous acid is a highly active oxidising agent, it is known to deactivate transport proteins,[41–43] destroy cellular cytochromes[44] and inhibit the replication of DNA.[45] Chlorine disinfection is also widely used to provide a residual chlorine level within distribution systems preventing/limiting the regrowth of microorganisms, as well as being used to backflush sand and GAC filters.

The issue with using chlorine disinfection so prolifically, is its ability to react with the organic matter present in the water, creating disinfection by-products (DBPs). Many of the by-products identified, as well as a mixture of them, have been shown to be cytotoxic, genotoxic and carcinogenic.[46] It is difficult to study all DBPs produced due to the complexity and diversity of their precursor NOM.

A prominent group of DBPs are a class of compounds known as trihalomethanes (THMs), a product of the haloform reaction. These were the first DBPs to be noted from the reaction between NOM and chlorine in drinking water production.[47,48] As such the EU, through the drinking water directive of 1998 set a standard limit of 100 μg/L for total THM concentration, this included chloroform, bromoform, bromodichloroform and dibromochloroform concentrations. THMs remain the only group of DBPs that are regulated in the EU, yet there are over 600 DBPs which have been identified, e.g haloacetic acids (HAAs), halonitriles and nitrosamines.[46,49]

As NOM is highly complex, there exists multiple reaction pathways that lead to the production of DBPs. These include oxidation, electrophilic substitution and electrophilic addition, only the first of these does not produce chlorinated by-products.[50] Figure 1.2 describes a possible mechanism by which chlorine, or more accurately hypochlorous acid, reacts with aldehydes and ketones to produce the main THM, chloroform.[50]

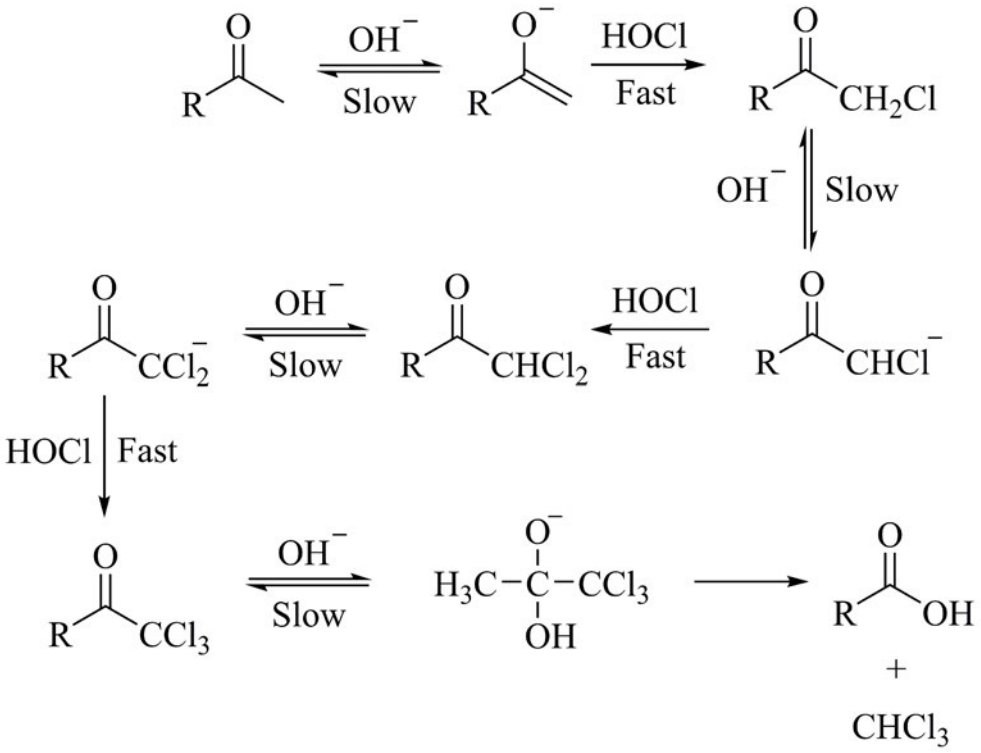


Figure 1.2: A proposed mechanism for the reaction between chlorine and aldehydes/ketones.[50]

Due to increasing NOM levels in surface waters, the THM limit is often breached, pushing potable water suppliers to find alternative methods of disinfection. One such method is using chlorine dioxide (ClO_2), which has been shown to inactivate bacteria by disrupting the cellular membrane and preventing protein synthesis.[51] This disinfectant is not widely used, despite being shown that it produces significantly less THMs and HAAs.[51,52] Chlorine dioxide is an unstable gas that has to be produced on site. There is also uncertainty regarding its effectiveness against *Legionella*,[53] the use of this disinfectant also results in the formation of chlorite (ClO_2^-) and chlorate (ClO_3^-), inorganic DBPs linked to changes in red blood cells and thyroid histology.[54]

A popular alternative treatment to chlorination is chloramination, a treatment which will be further discussed in Chapter 6. This technique involves the generation of chlorinated amines by the addition of chlorine and ammonia. The resulting chloramines are usually not used as a primary disinfectant they are weaker than free chlorine. The ammonia is usually added at the end of the disinfection stage, reacting with HOCl produced by chlorination as shown in Equation 1.2.

$$\begin{aligned} H_2O + NH_3 &\rightleftharpoons OH^- + NH_4^+ \\ NH_4^+ + HOCl &\rightarrow NH_2Cl + H_2O + H^+ \\ NH_2Cl + HOCl &\rightarrow NHCl_2 + H_2O \\ NHCl_2 + HOCl &\rightarrow NCl_3 + H_2O \end{aligned} \tag{1.2}$$

Chloramines are very stable, meaning they can provide a long lasting residual within distribution systems, reducing the likelihood of bacterial regrowth. Chloramines tend to be much weaker oxidising agents than hypochlorous acid, but have still been shown to produce THMs, although at much lower concentrations.[55] Chloramination does however result in the formation of new DBPs, many of which are nitrogenous including the potent carcinogen N-nitrosodimethylamine (NDMA).[56]

1.2.5 *Advanced Oxidation Processes*

Advanced Oxidation Processes (AOPs) are techniques developed for the oxidation of organic material within the water, the focus usually being to remove them prior to disinfection so that the level of DBPs produced is lower. These processes rely on the production of highly reactive radicals, usually the hydroxyl radical ($\cdot OH$), an extremely powerful oxidiser shown to be capable of degrading a wide range of organic species.[57]

Initial reactions between the hydroxyl radical and an organic species can involve the addition of the radical across double bonds or abstraction of a proton yielding water and an organic radical, the mechanism is dependant on the organic species present. Attack from a hydroxyl radical often leads to a cascade of oxidative radical reactions, ultimately leading to the mineralisation of the organic species.

There are two main classes of techniques for the production of hydroxyl radicals, photochemical and non-photochemical. Of the non-photochemical methods, most require the use of ozone (O_3) either at high pH or in conjunction with other chemicals such as hydrogen peroxide (H_2O_2). The decomposition rate of ozone in water increases with pH, due to the reaction with hydroxide ions, leading to the formation of the super-oxide anion radical ($\cdot O_2^-$) and the hydroperoxyl radical ($\cdot O_2H$), the super-oxide anion then reacts with remaining ozone to produce the ozonide radical ($\cdot O_3^-$) which is highly unstable and immediately decomposes to hydroxyl radicals, a summary is shown in Equation 1.3.

$$3O_3 + OH^- + H^+ \rightarrow 2(\cdot OH) + 4O_2 \tag{1.3}$$

Ozone can also be used to create hydroxyl radicals in waters with a neutral or acidic pH with the addition of an initiator, hydrogen peroxide. In water the hydrogen peroxide dissociates to the hydroperoxide ion (HO_2^-) which reacts with ozone to produce the ozonide radical and hydroperoxyl radical as per Equation 1.4.

$$O_3 + HO_2^- \rightarrow \cdot O_3^- + \cdot O_2H \tag{1.4}$$

Ozonation has been shown to result in the removal of NOM, with total organic carbon (TOC) reduction being consistently improved with an increasing addition of ozone.[58] It has also been shown to degrade several persistent organic contaminants that are commonly found in water, however there are concerns regarding the toxicity of partially oxidised pharmaceuticals. This along with the energy investment required for ozone production, is a significant drawback of ozonation.[59,60]

A well-known method of producing hydroxyl radicals without the need for ozone, is the Fenton process, a reaction reported in 1894,[61] involving the catalytic degradation of hydrogen peroxide using iron(II) as the catalyst, see Equation 1.5.

$$\begin{aligned} Fe^{2+} + H_2O_2 &\rightarrow Fe^{3+} + \cdot OH + OH^- \\ Fe^{3+} + H_2O_2 &\rightarrow Fe^{2+} + \cdot O_2H + H^+ \\ 2H_2O_2 &\rightarrow \cdot OH + \cdot O_2H + H_2O \end{aligned} \tag{1.5}$$

Initially iron(II) is oxidised to iron(III) forming the hydoxyl radical and hydroxide ion. The iron(III) can then be reduced back to iron(II) forming the hydroperoxyl radical in the process. This process relies on both iron(II) and iron(III) being present in solution meaning the rate is often dictated by the pH of the water, requiring highly acidic conditions. Another significant drawback to the Fenton reaction is the production of iron(III) oxide-hydroxide ($Fe(OH)_3$), which precipitates at pH values greater than 4, creating secondary pollution issues and loss of catalytic iron leading to reduction of reaction rates.[62,63] To try and offset these disadvantages there has been significant research into the use of heterogeneous iron. Its use eliminates the production of iron precipitates, the rate loss associated with this, and opens the technique up to a much wider pH range.[64–66]

The Fenton process has been implemented in the degradation of many pharmaceuticals and persistent organic species such as pesticides, with results showing $>$ 85% removal of selected endocrine disruptors.[67,68] The Fenton process has been extensively studied as a technique for the removal of NOM during potable water production, with some studies reporting $>$ 90% TOC removal.[69–71]

Almost all of the non-photochemical techniques can be enhanced with the introduction of UV light in the photochemical processes. Ozone undergoes photolysis under UV irradiation at 254 nm producing hydrogen peroxide as an intermediate, the hydrogen peroxide then also undergoes photolysis to generate hydroxyl radicals. Although it has been reported that when there are optically active compounds present in the water, the level of enhancement provided from the UV irradiation is reduced due to competitive absorption.[72] Organic structures shown to degrade due to either the O_3/UV or H_2O_2/UV processes include phenolic compounds (phenol, *p*-cresol)[72], small organic acids (oxalic acid, formic acid)[73,74], chlorophenols[75] and other chlorinated molecules.[76]

The Fenton process can also be enhanced with the introduction of UV-vis radiation ($<$ 600 nm) by the introduction of two possible reactions. The first being the photo-reduction of iron(III) to iron(II) ($<$ 580 nm) and if the light is of a short enough wavelength ($<$ 310 nm) peroxide photolysis as mentioned before; see Equation 1.6.

$$\mathrm{Fe(OH)}^{2+} + h\nu \rightarrow \mathrm{Fe}^{2+} + \cdot\mathrm{OH} \tag{1.6}$$

The light effectively increases the rate of iron(II) regeneration, increasing the oxidation rate of the reaction and reduces the amount of precipitate formed with respect to the non-photochemical Fenton process.[77,78]

Photocatalysis has been touted as a promising method for the oxidation of organic pollutants and NOM. This method utilises a semiconductor and UV/Vis light to produce hydroxyl radicals. A semiconductor is a solid material whose band gap is not excessively large (insulator) or too small (conductor), but electrons can still be excited with the use of photons of a reasonable energy (gap often given as less than 3eV).[79] The mechanism of photocatalysis is quite complicated, but the basic principles are as follows: The semiconductor is irradiated with a wavelength whose energy is equal or greater to the energy of the band gap, this causes excitation of electrons from the valence band to the conduction band, leaving holes (h^+) behind in the valence band. These electrons can dissipate the energy and return to the conduction band, losing heat in the process, a process known as recombination, which usually happens between 10 and 100 nanoseconds, in the bulk of the solid.[80] However, on the

surface of the catalyst the hole can react with adsorbed water to produce the hydroxyl radical and the excited electron can also react, usually with oxygen present in the solution, also often added to decrease recombination rate, as seen in Equation 1.7.

$$\begin{aligned} h^+ + \mathrm{OH}^-_{(adsorbed)} &\rightarrow \cdot\mathrm{OH}_{(adsorbed)} \\ \mathrm{O_2} + \mathrm{e}^- &\rightarrow \cdot\mathrm{O_2^-} \\ \cdot\mathrm{O_2^-} + \mathrm{H}^+ &\rightarrow \cdot\mathrm{O_2H} \\ 2(\cdot\mathrm{O_2H}) &\rightarrow \mathrm{H_2O_2} \end{aligned} \tag{1.7}$$

Titanium dioxide (TiO_2) is a very popular photocatalyst due to its high stability, low production costs, safety to both humans and the environment whilst showing reasonably efficient photoactivity. TiO_2 photocatalysis has been shown to result in the degradation of many organic pollutants such as chlorophenols[81], organic dyes[82,83] and compounds typical of those found in NOM.[84]

1.3 NUCLEAR MAGNETIC RESONANCE

Nuclear Magnetic Resonance or NMR, is a spectroscopic technique which uses magnetic fields to manipulate nuclei (hence the name), the information is obtained by observing the absorption and emission of electromagnetic radiation, in this case radiofrequency.

NMR is capable of not only providing a vast array of structural information but it can also afford compound concentrations and allow the study of molecular dynamics, whilst being a non-destructive technique. It is due to the aforementioned reasons that NMR has established itself as one of, if not the, most powerful techniques in analytical chemistry.

1.3.1 *Spin Angular Momentum*

All nuclei possess a quantum property known as spin, defined by the quantum number, I. The spin of a nucleus is a value greater than or equal to 0 in increments of $\frac{1}{2}$ and can be inferred by the rules presented in Table 1.4.

Table 1.4: Rules for the prediction of spin (I) from the protons and neutrons within a nucleus. This Table was adapted from reference[85].

Number of Neutrons	Number of Protons	Spin
even	even	0
even	odd	$\frac{1}{2}$ or $\frac{3}{2}$ etc
odd	odd	1 or 2 etc
odd	even	$\frac{1}{2}$ or $\frac{3}{2}$ etc

Nuclear physics shows us that nuclei with spin also possess spin angular momentum (P), a property related to the magnetic moment (μ) of a nucleus, as seen in Equation 1.8. Where γ is the magnetogyric ratio of the nucleus, this can be thought of as a measure of how magnetic the nucleus is, i.e. the larger γ, the stronger the magnet.

$$\mu = \gamma P \tag{1.8}$$

Nuclei which do not possess spin (I=0), do not possess a magnetic moment and are not observable, unfortunately several of the highly abundant nuclei critical in organic chemistry are spin 0 such as ^{12}C, ^{16}O and ^{32}S. In order to observe these nuclei, one has to rely on less abundant isotopes, ^{13}C for example. Nuclei with a spin greater than $\frac{1}{2}$ are termed quadrupolar and are observable, although they are characterised by fast relaxation times creating issues which will be described later.

Those nuclei termed spin half (i.e. I = $\frac{1}{2}$) are the easiest and most commonly studied, examples include ^{1}H, ^{13}C and ^{15}N. When these nuclei are placed within an external magnetic field ($\mathbf{B}_0$), the magnetic moments align themselves relative to the field in discrete quantised energy levels. The number of levels is given by 2I + 1, therefore, for spin half nuclei, there exists two energy levels, E_α and E_β, as shown in Equation 1.9, where $\hbar$ is the reduced Planck's constant.

$$E_\alpha = -\frac{1}{2}\hbar\gamma\mathbf{B}_0 \qquad E_\beta = +\frac{1}{2}\hbar\gamma\mathbf{B}_0 \tag{1.9}$$

1.3.2 *Larmor Frequency and the Vector Model*

When magnetic nuclei are subjected to an external magnetic field ($\mathbf{B}_0$), they precess about the orientation of that field in what is known as the Larmor

precession. The rate of this precession is known as the Larmor frequency ν, and is directly related to the strength of the external field and the magnetogyric ratio of the nucleus. (Equation 1.10)

$$\nu = \frac{-\gamma \mathbf{B}_0}{2\pi} \qquad \text{in Hz} \tag{1.10}$$

As with other forms of spectroscopy, NMR involves monitoring transitions between energy levels and this energy is provided in the form of electromagnetic radiation. The two states, E_α and E_β (seen in Equation 1.9) are aligned with and against the magnetic field respectively and are therefore of different energies, with the difference given below in Equation 1.11. This shows that the frequency of electromagnetic radiation needed to induce transitions between the two energy states is equivalent to the Larmor frequency.

$$\Delta E = h\nu = \frac{h\gamma \mathbf{B}_0}{2\pi} \tag{1.11}$$

^{1}H Larmor frequencies are typically hundreds of MHz, depending on the strength of ($\mathbf{B}_0$). With the ^{1}H frequency increasing by 42.58 MHz per Tesla, NMR frequencies fall within the radiofrequency portion of the electromagnetic spectrum.

As E_α is aligned with the magnetic field, it is lower in energy than that of E_β. With all energy differences there exists a population difference, as described by the Boltzmann distribution 1.12. N_α and N_β represent the number of nuclei in that orientation, k_B is the Boltzmann constant and T is temperature.

$$\frac{N_\alpha}{N_\beta} = \exp(\frac{\Delta E}{k_B T}) \tag{1.12}$$

Based on Equation 1.11, the energy difference between both states is extremely small, and therefore the population difference is also incredibly small. It is for this reason that NMR is insensitive in comparison to other methods of spectroscopy. Figure 1.3, shows how this population difference can be represented as a bulk magnetisation vector (M), which at equilibrium is stationary and lies along the z axis in the vector model.

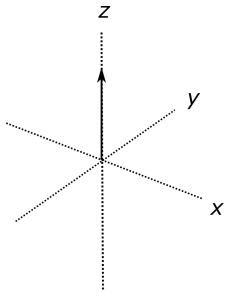


Figure 1.3: Vector model representation of bulk magnetisation.

The vector model is a method of visualising how the bulk magnetisation is manipulated throughout an NMR experiment and has two iterations, the first being the laboratory frame, where the x, y and z axes are fixed in space.

The second and arguably more useful, is known as the rotating frame. When considering a single nucleus environment, the rotating frame can be pictured as having the x and y axes rotating at the Larmor frequency. During most experiments there are often multiple nuclear environments, the rotating frame in this case is considered to be rotating at the carrier frequency for a given experiment.

The vector diagram is an extremely useful way of visualising where the M will be during pulse sequences. Figure 1.4 shows examples of how different pulses manipulate the initial z magnetisation (M_0).

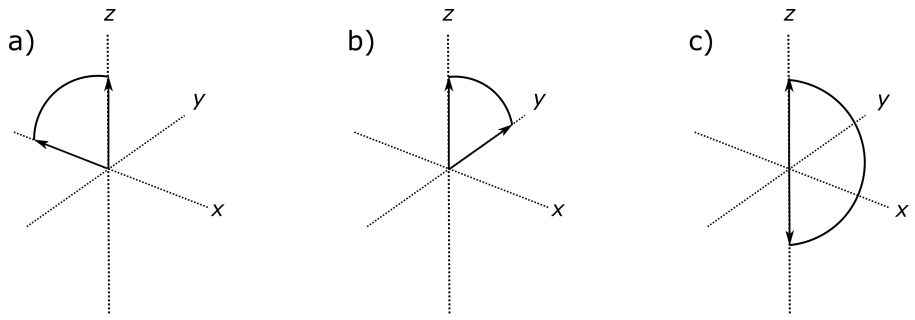


Figure 1.4: Vector model representation of a) a 90° pulse from the y axis b) a 90° pulse from the x axis and c) a 180° pulse from the -y axis.

As previously mentioned, when the spins are in the presence of a magnetic field, they precess about the Larmor frequency. When the spins have been manipulated via radio frequency pulses away from z and into the transverse plane (x, y plane), they continue this precession. Modern spectrometers utilise a detector within the transverse plane and as the spins precess across this detector, they create a time dependent oscillating signal, which can be converted into the frequency domain. For a single frequency this can be easily determined from the time domain. However, in a realistic situation, there are usually

many more frequencies, making extraction of these difficult/impossible, so this deconvolution of frequencies is almost exclusively performed using a mathematical procedure known as the Fourier transformation (see 1.13).

$$f(\omega) = \int_{-\infty}^{\infty} f(t) \mathrm{e}^{i\omega t} dt \tag{1.13}$$

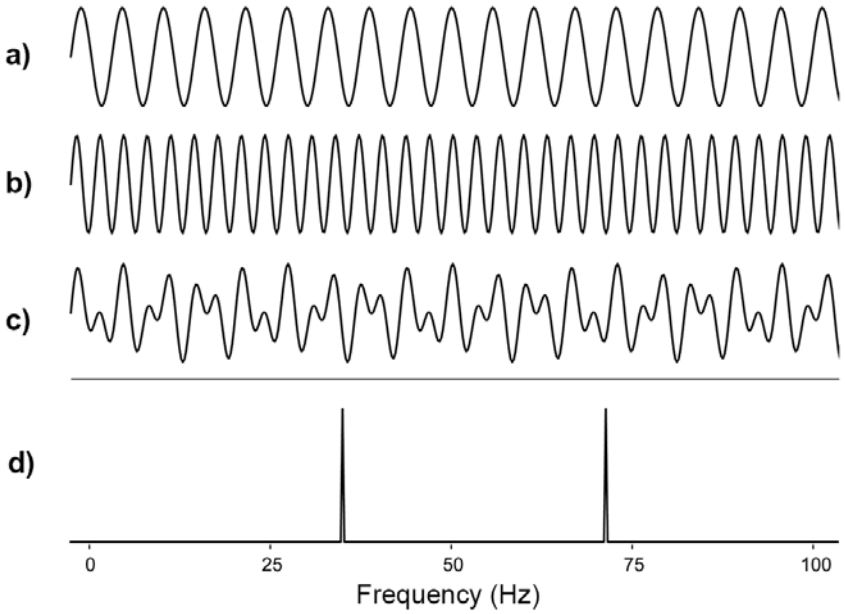


Figure 1.5: Example of the Fourier transform. **a)** a 35 Hz sine wave, **b)** a 70 Hz sine wave, **c)** the sum of a and b, **d)** the product of Fourier transformation on c.

1.3.3 *Relaxation*

As mentioned previously, when spins are subjected to an external magnetic field, there exists a population difference between 2I + 1 quantised energy levels. This population difference is represented by the bulk magnetisation being in the $+z$ axis of the vector diagram (Figure 1.3). This is the equilibrium state of the system and when we manipulate spins using radiofrequency (RF) pulses (input energy into system) the population difference is disturbed. Over time the system loses this increased energy through relaxation processes and re-establishes the equilibrium state.

There exists two main relaxation processes, T_1 (longitudinal relaxation) and T_2 (transverse relaxation). T_1 relaxation corresponds to the recovery of magnetisation in the $+z$ axis (longitudinal component). In order to achieve this, the system had to lose energy and it does this by transferring it into the

surrounding bulk via heat. This process follows an exponential function and for ^{1}H nuclei tends to be on the order of seconds (1 - 5s) and on the order of seconds to minutes for ^{13}C.

These relaxation times are large due to spin relaxation not being a spontaneous process (rate is ≈ 0) and there being no efficient pathway for the transfer of this energy. The relaxation occurs mainly via stimulation by a suitable fluctuating field (at or near the Larmor frequency). These fields arise from various sources, dipole-dipole interactions for example, and the time dependent fluctuation of the field is induced mainly due to molecular tumbling in solution. As this tumbling is chaotic, there exists a multitude of frequencies as molecules will collide with one another, however the tumbling of molecules is characterised by the average time it takes for the molecule to rotate through one radian, the rotational correlation time (τ_c). The shorter τ_c, the more rapid the tumbling.

In solution spectroscopy at or near room temperatures the rotational correlation time is usually several orders of magnitude higher than the Larmor frequency (ν) making relaxation inefficient and leading to longer T_1 times. An increase in solution viscosity (lower temperatures) or an increase in molecular size, results in an increase in τ_c, bringing the field fluctuations closer to ν and enhancing relaxation, i.e. T_1 decreases. There comes a point where the increase in molecular size or solution viscosity results in these field fluctuations becoming slower than ν and as a result the relaxation efficiency will also decrease, resulting in a longer T_1.

Practically, a longer T_1 results in a longer experiment time, for quantitative results it is recommended to wait 5 x T_1 in order to ensure that 99% of the magnetisation has returned to z before the next scan.

Transverse relaxation, T_2, rather than being the return of the bulk magnetisation vector to the z axis, is the dephasing of magnetisation within the transverse plane. Immediately following a 90° y pulse, the bulk magnetisation vector is along the x axis. In the rotating frame and assuming we are on resonance with the spin, one would expect the bulk magnetisation vector to remain along the x axis and remain in phase. However in reality, identical spins throughout the sample don't experience the same magnetic field all of the time. Fluctuating magnetic fields caused by molecular tumbling act as random RF pulses along the z axis and cause some of these *identical* spins to change their phase. The result of this is a spreading of the spins in the transverse plane, ultimately leading to a loss of net magnetisation as seen in Figure 1.6, again following exponential decay. As T_2 relaxation describes loss of magnetisation in the transverse plane and T_1 relaxation describes the return of the magnetisation to z, T_2 can never be longer than T_1.

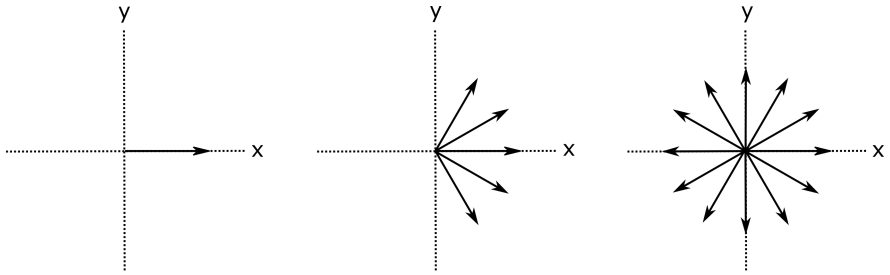


Figure 1.6: Vector model representation of spins spreading throughout the transverse plane during the T_2 relaxation process.

Unlike in T_1, where the most efficient relaxation occurs at the Larmor frequency, T_2 relaxation is caused by magnetic fields of any frequency. T_2 relaxation becomes more efficient at slower τ_c, therefore, increasing the molecular weight of the molecule or increasing the viscosity of the solution will result in a decrease in the T_2 relaxation time. T_2 relaxation times for low molecular weight species in non-viscous solvents are very close if not the same as the T_1 relaxation time.

Practically, T_2 relaxation times are important, as they are linked to the linewidth of a frequency (see Equation 1.14). A short T_2 relaxation time corresponds to a faster spreading of spins in the transverse plane, resulting in a broadening of the spectral line.

$$\Delta\ \nu_{\frac{1}{2}} = \frac{1}{\pi\ T_2^*} \tag{1.14}$$

where $\Delta\ \nu_{\frac{1}{2}}$ is the linewidth at half height of the signal, and T_2^* describes the overall relaxation time with contributions from both genuine relaxation processes, T_2, and magnetic field inhomogeneity (usually denoted $T_{2(\Delta \mathbf{B}_0)}$)

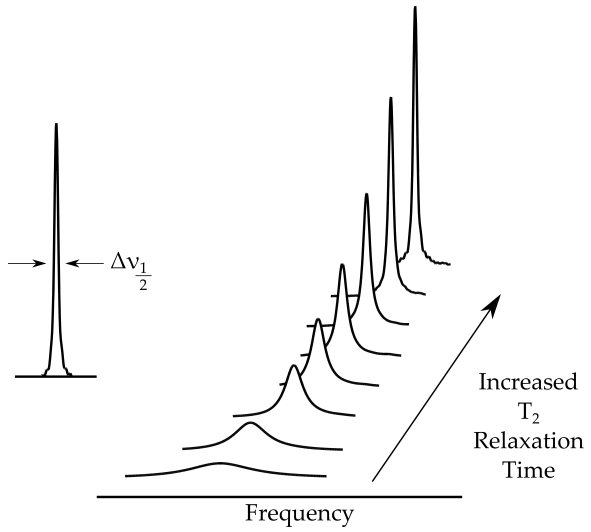


Figure 1.7: Selection of frequencies highlighting the effect on linewidth with changing T_2 relaxation times

It was mentioned earlier that the NMR spectrometer detects an oscillating signal as the spins precess throughout the transverse plane, because T_2 relaxation corresponds to an exponential loss of coherent signal in the transverse plane, the observable NMR signal known as a free induction decay or FID, also decays via an exponential function as shown in Figure 1.8

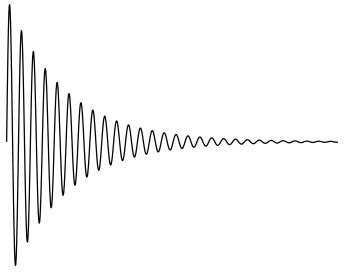


Figure 1.8: FID showing an exponential decay as a result of T_2 relaxation

1.3.4 *Chemical Shift and Coupling*

With regards to structural determination two of the most useful pieces of information obtainable via NMR have to be the chemical shift (δ in ppm) and coupling constants (J in Hz). Using protons as an example, in a 14.1 Tesla magnet, protons within a molecule will resonate at 600 MHz, however, in reality these protons have slightly different precession rates dependant on their local environment. This is known as their chemical shift and is mainly caused by electrons within the system. Interacting with the external magnetic fields, electrons produce induced magnetic fields that act to augment the strength of

the external magnetic field ($\mathbf{B}_0$) to varying extent; this is known as shielding and deshielding effects and depends on the local electron density.

As frequencies are dependant on the strength of $\mathbf{B}_0$, they are normalised into the units of ppm using Equation 1.15. This equation means that 1 ppm is equal to the Larmor frequency of a nucleus divided by 10^6, i.e. in a 600 MHz spectrometer for proton, 1 ppm constitutes 600 Hz.

$$\delta = 10^6 \ (\frac{\nu - \nu_{ref}}{\nu_{ref}}) \tag{1.15}$$

Chemical shift values for proton cover a reasonably small range 0 - 12 ppm, however other nuclei such as ^{19}F are much more sensitive to their local environment and cover a much larger range of >350 ppm. Protons on non-electron withdrawing groups will be relatively shielded and therefore have a corresponding low chemical shift, methyl groups for example have a chemical shift of around 0.5-2 ppm, however if the methyl group is directly attached to an electronegative atom such as an oxygen (methoxy group), it becomes relatively deshielded with a chemical shift of 3-4 ppm.

Valuable structural information can also be obtained through spin-spin coupling. Let's consider two different proton environments adjacent to one another, H_A and H_B. From the point of view of H_A, H_B can either be aligned with or against $\mathbf{B}_0$, adding or subtracting respectively to the strength of magnetic field experienced by H_A. This changes its resonant frequency by $\pm\frac{J}{2}$, where J is known as the coupling constant, the distance in Hz that the resonant frequency is split by. This relationship is mutual, so both H_A and H_B would be split by the same J value.

Coupling is a phenomenon that is transmitted through bonds, protons will usually exhibit coupling to appropriate nuclei between 2-4 bonds away. Typical ^{1}H - ^{1}H J values range between 1 - 20 Hz whereas ^{1}H - ^{13}C values range between 1 - 20 Hz for ^{2-4}J (2-4 bonds) but are much larger for ^1J between 120 - 160 Hz. The values of these couplings can be extremely useful in determining molecular structures, particularly in small molecules, as will be discussed later in this thesis.

Where J-coupling requires a pair of spins to be close bond wise, the Nuclear Overhauser Effect (NOE) is a through space interaction, allowing the user to gain information on the 3D geometry of the molecule. The NOE is a very weak effect resulting from cross relaxation via dipolar interactions, the size of this NOE is influenced by the relaxation rate, which is proportional to r^{-6}, where r is the internuclear separation. The distance limitation for ^{1}H is roughly 5 Å, meaning both inter and intramolecular correlations can be observed.

1.3.5 *Higher Dimensional NMR*

In 1D NMR experiments, multiple nuclei may be irradiated throughout the pulse sequence and acquisition, but only one frequency dimension is detected during acquisition. i.e. a 1D ^{13}C experiment with ^{1}H decoupling, or 1D ^{13}C DEPT experiment. In a 2D experiment, two frequency dimensions are recorded, although the second dimension could be the same as the first (eg in ^{1}H - ^{1}H COSY), it could also be a diffusion dimension. 2D experiments are used to map interactions within (or between) molecules in the sample, interactions such as scalar coupling (J) or dipolar couplings (NOE) can be mapped.

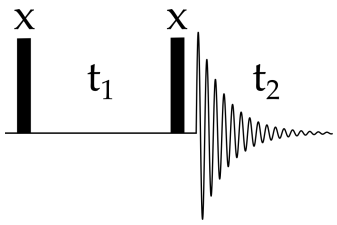


Figure 1.9: Basic 2D COSY pulse sequence.

Generally, in order to create a frequency dimension, it is necessary to regularly sample the magnetisation as it varies in time. One of these time/frequency domains originates from sampling the FID during acquisition, the other then, must be generated during the pulse sequence. This will be described using the basic COSY sequence presented in Figure 1.9.

The first 90° pulse puts our magnetisation into the $+y$ axis, after which it will evolve (begin precessing) according to its chemical shift offset as a function of t_1. The second 90° pulse moves the y magnetisation component to -z, but the x component is unaffected and the signal proceeds to t_2. If t_1 is set to zero the two 90° pulses place the magnetisation in -z, there is no transverse magnetisation and hence no signal. As the t_1 interval is increased from zero, the x component increases, with a maximum being at 90°, beyond this the x component decreases and then becomes negative, i.e. creates a sinusoidal wave. This signal is also subject to relaxation during the t_1 interval, resulting in a gradually generated FID known as an interferogram and is said to have been generated indirectly. In a situation where we don't observe any ^{1}H - ^{1}H coupling between protons ν_A and ν_B, the Fourier transformation of the direct and indirect dimensions will result in a spectrum like in Figure 1.10a, where the individual frequencies for ν_A and ν_B lie on the diagonal, i.e. they are self-correlated. This shows no additional information than a simple 1D ^{1}H experiment would, however if the two protons are coupled to one another, part of the magnetisation is transferred from one spin to the other during the final 90° pulse, this is known as coherence

transfer. As only part of the magnetisation is transferred to the other frequency, at the second detection period there will exist a peak at the coordinates ν_A - ν_A, like in the case without coupling, but also more importantly an off diagonal cross peak at ν_A - ν_B, as shown in Figure 1.10b.

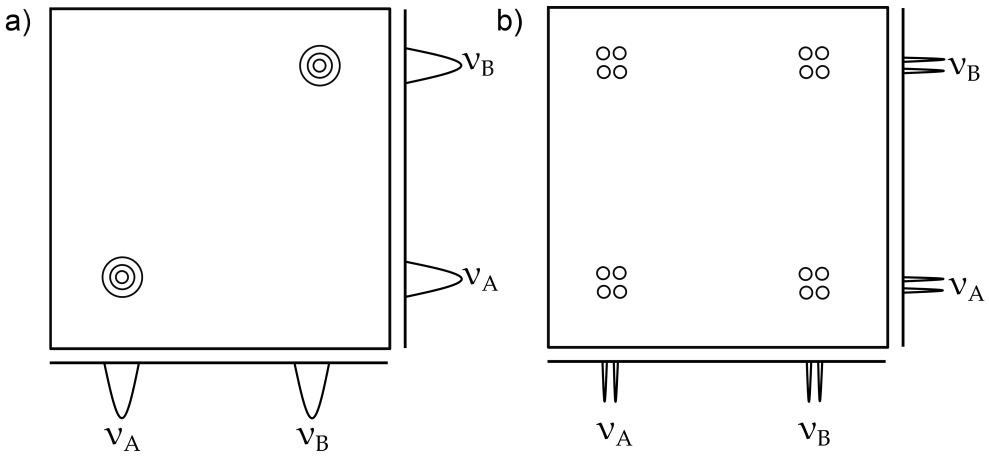


Figure 1.10: Example ^{1}H - ^{1}H COSY spectrum where ν_A and ν_B are a) uncoupled and b) J coupled.

This technique of indirect chemical shift sampling has been utilised to create an extensive set of NMR experiments with two or more dimensions, examples of the most common are shown in Figure 1.11. With the addition of each dimension, one must add an additional incremental delay, i.e. a 3D experiment will have a t_1, t_2 and a t_3 period. As each additional dimension must be sampled appropriately, 3D experiments require longer overall experimental times. In practice these are sampled with fewer time domain points and a compromise has to be achieved between resolution and overall experiment time. Signal aliasing or using non-uniform sampling can help this issue.[86,87]

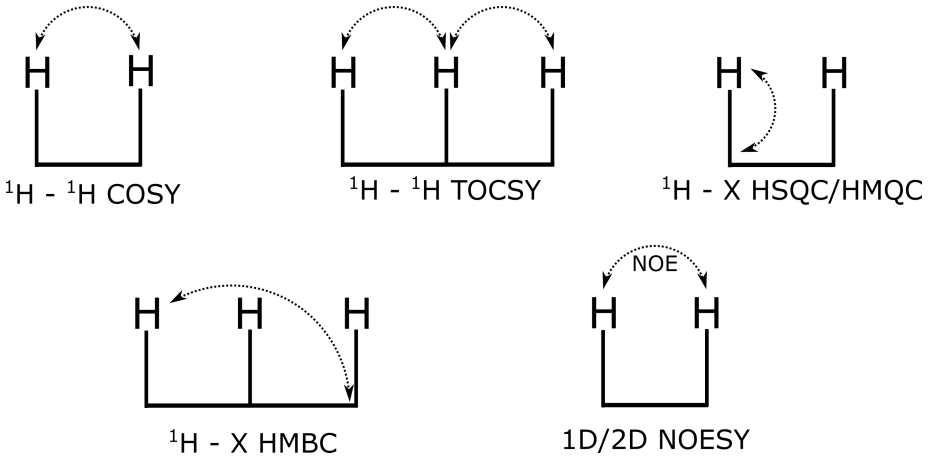


Figure 1.11: Schematic showing the structural information a user can obtain when using some of the most common NMR experiments.

Figure 1.11 highlights some of the most common 2D experiments, the first being a COSY (Correlation Spectroscopy) experiment. A homonuclear ^{1}H - ^{1}H COSY is shown, where magnetisation is transferred between J-coupled protons, usually separated by 2-3 bonds. The Total Correlation Spectroscopy experiment (TOCSY) is similar to the COSY in that magnetisation is transferred via J-coupled protons, but here there is a mixing period (spin-lock) which allows the magnetisation transfer to be extended beyond adjacent protons. The result of the TOCSY is that correlations within entire spin systems can be observed. A short mixing time results in magnetisation only being transferred through large couplings (i.e. usually closer protons) and fewer bonds, so by manipulating the mixing time it is possible to determine the sequence of connectivity within the spin system.

Experiments can also be heteronuclear, transferring magnetisation between two different nuclei, i.e. ^{1}H to ^{13}C. An example of this is the ^{1}H - ^{13}C Heternonuclear Single Quantum Coherence experiment (HSQC), a cornerstone of organic structure determination. This experiment provides ^{1}H - ^{13}C single-bond coupling information and uses the Insensitive Nuclei Enhanced by Polarisation Transfer (INEPT) method of magnetisation transfer. This element boosts experiment sensitivity by transferring polarisation from the more sensitive nucleus (^{1}H) to the lower sensitivity nucleus (^{13}C) prior to the evolution period (t_1). The magnetisation is then transferred back to proton for detection during t_2. Heteronuclear experiments like this are reduced in sensitivity relative to the homonuclear experiments due to low natural abundance of the X nucleus, in this case ^{13}C (1.107 % natural abundance).

The Heternonuclear Multiple Bond Coherence (HMBC) experiment allows observation of long range ^{1}H - ^{13}C correlations. These make it easier to piece

together molecular fragments with low proton density. This sequence takes advantage of the large difference between ^{1}J (large) and ^{n}J (order of magnitude smaller), often by employing J-filters, removing (or reducing) interference from the one-bond cross peaks. A particularly useful feature of the HMBC experiment is that correlations can be obtained across heteroatoms, such as an oxygen (ether) or nitrogen (amide) moiety, allowing for the study of a wide variety of molecules.

1.3.6 *Uses of NMR in the Study of NOM*

NMR is a highly reproducible, non-invasive technique with a vast library of experiments to draw from that allows the user to provide structural information on unknown compounds without the need for prior knowledge. It is for these reasons that NMR spectroscopy is increasingly being used to answer questions related to environmental samples.

A major restriction on the NMR analysis of NOM is the concentration of individual compounds. As NOM consists of thousands (possibly tens of thousands) of compounds, at environmental DOM concentrations (1-20 mg/L), each individual compound is going to be present at very low levels. To overcome this, the majority of applications involve the use of the more sensitive homonuclear experiments such as 1D ^{1}H experiments. Throughout this thesis a classification of the 1D proton spectrum was used, adapted from Simpson *et al*,[88] to separate classes of resonances into likely chemical species as shown below in Figure 1.12. The integrals obtained from these regions can then be normalised and compared across spectra/samples.

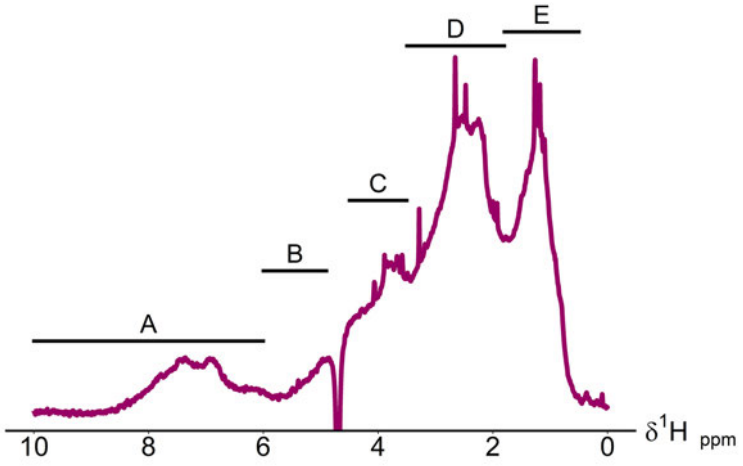


Figure 1.12: An example NOM ^{1}H spectrum, acquired using a NOM standard. Labelled regions correspond to a) aromatics, b) unsaturated resonances, c) carbohydrates, d) Carboxyl Rich Alicyclic Molecules (CRAM) and e) aliphatics.

In order to overcome the lack of resolution provided by 1D ^{1}H NMR, many use nuclei with a larger chemical shift dispersion like ^{13}C, although for NOM samples there is still significant local overlap, highlighting that the sample complexity is still an issue.[89] Sensitivity permitting, one can find more resolution through the inclusion of other dimensions, further separating chemical constituents of NOM, uncovering some of the information lost in 1D spectra.

The first example of a 2D NMR technique being applied to NOM was the *J*-resolved experiment, an experiment which has ^{1}H chemical shifts in one dimension, ^{1}H J constants. The authors of this study used this 2D experiment along with some 1D spectral editing techniques to elucidate partial structures of NOM components.[90]

For a long period of time, many thought that NOM consisted of large molecular weight species, with average molecular weights in the mass range of $>$ 50 kDa. André Simpson used NMR, specifically diffusion spectroscopy (DOSY) to shed more light on this phenomenon. At high concentrations, the molecules within NOM display diffusion constants similar to that of large macromolecules ($>$66 kDa), but at low concentrations, it becomes apparent that this is artificial and aggregation is common place in these complex matrices. DOSY experiments implied that NOM is composed of lignins, carbohydrates and peptides with molecular weights as low as 200 Da. This however, does not mean that there are no macromolecules present within NOM, the DOSY experiments also showed the presence of polysaccharides or proteins in an agricultural NOM sample.[88]

1.4 FOURIER TRANSFORM ION-CYCLOTRON RESONANCE MASS SPECTROMETRY (FT-ICR-MS) THEORY

Of all the mass spectrometry detection methods, those that apply Fourier transformation are capable of providing the highest resolution. The three main classes of FTMS include the Orbitrap, the ion trap and the technique featured here, Ion-Cyclotron Resonance (ICR). FT-ICR-MS, of these three, delivers the highest resolution and therefore the greatest mass accuracy. It is for this reason that it is the MS technique of choice with regards to the study of complex mixtures such as NOM.

1.4.1 *FT-ICR-MS Instrumentation*

A schematic of a typical FT-ICR is shown in Figure 1.13. The path an ion takes through the spectrometer can be divided into 4 major sections. First the molecules within the sample must be ionised. This is usually performed using an Atmospheric Pressure Ionisation (API) source such as Electrospray Ionisation (ESI)(**1** in Figure 1.13), discussed in more detail in the next subsection.

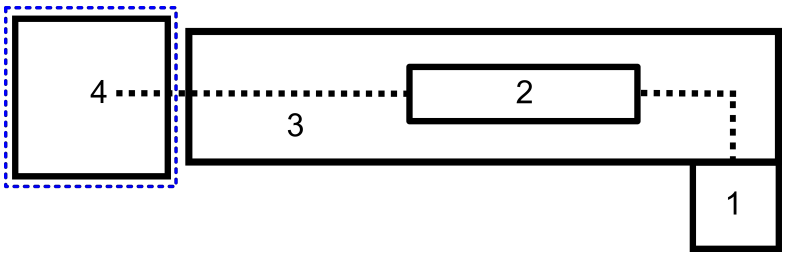


Figure 1.13: Simplified schematic of a Bruker SolariX FTICR instrument. **1** API source, **2** quadrupole and a hexapole collision cell, **3** transfer optics and multipole, **4** ICR cell. The black dashed line indicates the path of ions, whilst the blue dashed line indicates the magnetic field.

Ions are then funnelled into an initial mass analyser, a quadrupole, consisting of four rods set parallel to one another, these rods apply electromagnetic fields which deflect ions of given masses. Here the user can set mass limits for passing ions, a lower and an upper limit, effectively allowing for the filtration of a specific mass range. From the quadrupole, ions are then accumulated within a hexapole known as a collision cell, here ions can be accumulated for a given length of time, allowing the user to build up enough ions to optimise signal to noise; there are downsides to having too long an accumulation time which are described later. Within the collision cell, ions can also be fragmented using

Collision Induced Dissociation (CID) a technique which involves applying a collision voltage and pumping in a neutral gas. Here, molecules collide together and fragment, hence why this part of the spectrometer is given the name of collision cell despite this not being its only use.

After leaving the accumulation cell, ions are transferred through a multipole (**3**) to the ICR cell situated within an superconducting magnet (**4**). The time allowed for transfer to occur between the collision cell and ICR cell is known as the Time of Flight (TOF) and can be tuned by the user to act as a way of filtering masses. Heavier ions take longer to transfer between cells and can therefore be filtered out by shortening this time. For studies on NOM, a TOF between 500 and 800 ms is usually reported. As the ions travel throughout the system they are brought from atmospheric pressure (usually, but not always) down to the ultrahigh vacuum that FT-ICR requires, by the time they are in the cell, the pressures are in the order of 1 x 10^{-10} mbar.

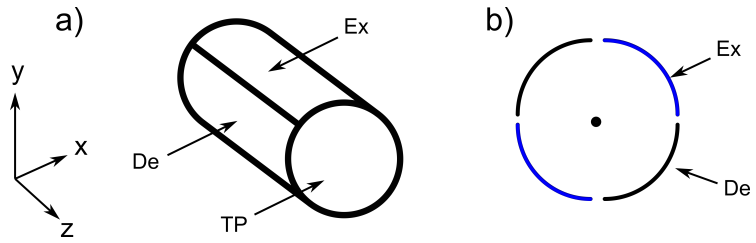


Figure 1.14: **a)** A simplified schematic of a cylindrical ICR cell showing the orientation of the detection plates (De), the excitation plates (Ex) and the trapping plates (TP) and **b)** The same cell viewed from along the z axis.

A simplified schematic of the cell is shown in Figure 1.14, this diagram is based on the 'infinity' cell, common to the Bruker SolariX system. This cell has a pair of trapping plates, across which a voltage is applied in order to contain the ions. The cell also has two excitation and two detection plates that are responsible for optimisation of the ion distance to the detection plates and then signal detection, respectively. Throughout this PhD the "infinity" cell was replaced with what is known as the ParaCell or the dynamically harmonised cell, which is capable of producing more homogeneous electric fields leading to increased ion stability and allowing for greater mass accuracy.

1.4.2 *Fundamental Theory of ICR*

Ions that are trapped within the ICR cell, or any magnetic field, will experience a constant inward force perpendicular to that field known as the Lorentz force. This inward force develops an equilibrium with the centrifugal force as the ions

circulate around the centre of the magnetic field, causing ions to circulate at a frequency which is characteristic of their m/z value. This frequency is known as the cyclotron frequency and is shown in Equation 1.16, where the angular cyclotron frequency (ω in rad s^{-1}) is related to the elementary charge (q) and the magnetic field strength (B in Tesla). Since the magnetic field is fixed within an ICR spectrometer, the cyclotron frequency is directly related only to the mass and the charge of the ion.

$$\omega = 2\pi f = \frac{qB}{m} \tag{1.16}$$

Once in the cell and orbiting, the ions are very close to the centre of the magnetic field, i.e. relatively far from the detection plates. The application of a broadband RF pulse excites these ions to a larger radius closer to the detection plates. As Equation 1.16 shows, the cyclotron motion does not have a kinetic energy dependence, therefore when these ions are excited they retain their original characteristic frequency. Ions travel around this enlarged radius in a coherent manner, usually termed an "ion cloud". When they pass across the detection plate they induce a current, creating a sine wave transient, see Figure 1.15.

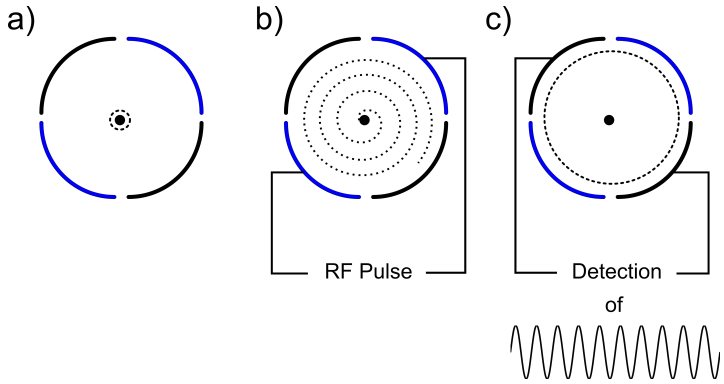


Figure 1.15: A highly simplified diagram showing motion of ions throughout the detection scheme. **a)** Ions are in orbit around the center of the magnetic field, **b)** these ions are then excited using an RF pulse to a larger radius and finally **c)** as ions travel across the detection plates they create a sine wave transient which is then Fourier transformed to yield the cyclotron frequency of the ion.

As FT-ICR is performed under ultra-high vacuum and in a perfect system, the ion cloud could continue orbiting at this enlarged radius indefinitely. However, nothing is perfect and there are several limitations to the transient lengths. One such limitation arises from ion-ion or ion-neutral gas (imperfect vacuum) collisions that reduce the number of ions within the ion cloud over time, often meaning that a longer transient length results in lower sensitivity. Another

limitation to transient length arises from motions other than the cyclotron motion, known as the magnetron motion and axial trapping motion (caused by imperfect trapping plates). The physics behind these motions is quite complex and on modern spectrometers, such as the one used throughout this thesis, these motions are orders of magnitude lower than the cyclotron motion and are not observed, therefore will not be described further. Practically for broadband observation of complex mixtures, transient lengths of less than 5 seconds are commonplace.

Other negative effects observed in the ICR cell include space charge effects and ion cloud instability.[91,92] Space charge effects are the result of electrostatic interactions between charged species trapped within a finite space. This ultimately results in errors in the measured frequency and therefore reduced mass accuracy.[91] Ion cloud instability can occur when the vacuum is inefficient and/or the total number of ions within the cell becomes too large. This can result in the same mass being observed at slightly different frequencies in alternating scans/transients, resulting in what is often termed 'split' peaks, where two close signals are observed for the same mass.

1.4.3 *Ionisation Methods*

There are several commercially available ionisation sources, with each method varying in ionisation mechanism, ultimately showing a bias towards a certain chemistry. Another way to think of this is that each technique has a unique analytical window into the sample.

The methods described in this section focus on the ionisation techniques/sources used in this thesis.

1.4.3.1 *Electrospray Ionisation*

By far the most commonly used ionisation source, Electrospray Ionisation (ESI) is a soft technique (limited fragmentation) which is utilised at atmospheric pressure.[93]

Generally, a dilute sample is injected by a syringe, at a low flow rate (1-20 μL/min) into a nebuliser through which a neutral gas (often nitrogen) flows at high pressure (2-3 bar). This neutral gas helps vaporise the sample solution into a fine mist of solvent droplets. Across the end of this nebuliser (outlet), a very high voltage is applied, usually between 2-6 kV, which induces a charge on these small solvent droplets. At this point the charged droplets are being pulled into the spectrometer vacuum, here they pass through a drying gas which

concentrates the charge further until the charge is no longer only on the solvent, but is passed on to the analyte of interest, producing ions.[94] ESI requires the presence of functional groups possessing labile protons, such as carboxylic acids or amines. It will not produce ions from non-polar species.

For NOM analysis ESI is usually performed in negative mode, in which ions are formed by the loss of a H^+ ion, meaning the parent ion (M) detected is a M^- ion. It is possible to utilise ESI in positive mode (commonly done), however this often results in the production of adducted ions. Rather than resulting in the loss of a proton like in negative mode, there is the addition of another ion, usually a proton or sodium to produce $[M+H]^+$ or $[M+Na]^+$ respectively. The inclusion of these adducts increases the complexity of spectra, therefore longer transient lengths are required for unambiguous assignment. For this reason, positive mode is not typically performed for complex mixture analysis. It should be noted that adduction can also occur in negative mode (although less common), usually in the form of chloride adducts, $[M+Cl]^-$, this is one reason complex mixture samples, particularly NOM, are often desalted as part of the sample preparation.

ESI is the ionisation technique that dominates the complex mixture analysis field. It is relatively easy to create a stable spray, meaning a stable introduction of ions into the spectrometer. Resulting in a consistent number of ions being introduced to the detection cell from transient to transient, allowing the user to tune the instrument for improved resolution.

1.4.3.2 *Laser Desorption Ionisation*

Laser Desorption Ionisation (LDI) is an ionisation technique performed on solid samples. The mechanism of ionisation is not fully understood, however, it is reported that the laser ablates material from the surface creating a microplasma of ions and neutral molecules near the sample surface.

The laser used on the instrument utilised throughout this thesis (Smartbeam-II, ND:YAG μ3 laser) produces photons at 355 nm, meaning molecules with conjugated systems are much more likely to ionise. Moreover, it is reported that the produced ions are most likely to be fragmentation products of the parent molecule if the parent is $\approx$ 500 Da or greater, preventing full characterisation of the parent molecule.

A major advance in the technique, known as matrix assisted LDI (MALDI), solved this issue of fragmentation by crystallising the analyte of interest with a small organic molecule which has strong absorption at the laser wavelength. Again, the ionisation mechanism is not fully understood, but MALDI is known to be more sensitive than most laser ionisation techniques.[95,96] It is not necessary

to tune the laser wavelength to the analyte of interest because it is the matrix which absorbs the laser pulse. This also means that the variety of analytes amenable to this technique is much greater than with LDI.

The reason MALDI is mentioned here, despite not being used, is that there is evidence that NOM samples, possessing abundant small aromatic molecules, can take the place of a conventional matrix and promote ionisation.

(MA)LDI however, is not commonly implemented in complex mixture analysis as it suffers from low shot to shot reproducibility, mainly caused by a heterogeneous surface.

1.4.4 *Resolution and Calibration of MS Spectra*

Resolving power refers to a peaks linewidth relative to its mass, as shown in Equation 1.17, where R is resolving power, m is mass and $\Delta m_{\ FWHM}$ is the width of the peak at a given height, here FWHM refers to 50% or Full Width Half Maximum.

$$\mathrm{R} = \frac{m}{\Delta m_{(FWHM)}} \tag{1.17}$$

An instrument with a greater resolving power is capable of separating peaks with very similar masses, i.e. increased resolution, something that is sorely needed when analysing complex mixtures. As previously mentioned, FT-ICR-MS is capable of substantially greater resolving power than most other MS techniques, with resolving powers beyond 10^6 routinely acquired. Other techniques, such as Orbitrap and TOF, routinely achieve resolutions of 10^5 and 10^4, respectively.

As mentioned earlier, the frequency of ions within the cell can be affected by phenomenon such as space charge effects, ion cloud stability and inhomogeneous electric fields. To overcome these issues, FTMS spectra need to be calibrated in order to obtain accurate masses. The frequency needs to be converted to mass/charge (m/z), this can be done using multiple equations, two of which are shown in Equations 1.18 and 1.19.[97,98]

$$\frac{m}{z} = \frac{A}{f} + \frac{B}{f^2} \tag{1.18}$$

$$\frac{m}{z} = \frac{A}{f} + \frac{B}{f^2} + \frac{C \times \mathrm{TIC}}{f^2} \tag{1.19}$$

In the above equations, A and B represent constants for magnetic and electric field coefficients respectively, these are determined experimentally using a standard sample of known mass, such as arginine clusters; TIC represents

the Total Ion Current and f is frequency. Some researchers, including Bruker developers also implement a third calibration coefficient, C. This initial calibration is likely to result in a spectrum with a reasonable mass accuracy (<10 ppm). It is however instrument specific, therefore if any instrumental parameters were to be changed, it's reliability is likely to be negatively affected. The instrument calibration is rarely changed as a much better calibration is achievable through the implementation of external and internal (re)calibrations performed post acquisition. For internal calibration a series of known species across the mass range of the spectrum are selected and a correction function is fitted to their mass error, this calibration method for complex mixtures can result in a spectrum with extremely high mass accuracy (<0.5 ppm). External calibration involves creating an internal calibration for a sample, creating that correction function and then applying this function to another spectrum. External calibration performs worse than an internal calibration.[99]

After calibration, the resulting masses can be assigned to molecular formulae, this is often done on the basis of accurate mass alone, as atoms have exact masses this is usually all that is required. By convention, ^{12}C has been assigned a mass of exactly 12 Da and is used as the basis for all other atoms, the majority of which have been assigned a highly accurate mass. Therefore in order to assign a molecular formula, all one needs to do is calculate what combination of elements will provide the experimentally measured mass, within a degree of error.

1.4.5 *Interpretation of FT-ICR-MS Spectra in the Study of NOM*

Throughout the study of complex mixtures, it has become apparent that many of the compounds form homologous series, i.e they are related to each other by a repeating unit. Particularly for NOM, this is likely due to the same base structure degrading to differing molecules that possess the same skeleton. One method of visualising these homologous series is known as the Kendrick mass defect (KMD) plot and involves modifying the mass of the suspected repeating unit. The most common repeating unit is the CH_2 unit which instead of being 14.01565 Da is now defined as 14, another common unit to use is CO_2, being 44 instead of 43.98983.

$$\mathrm{KM} = \mathrm{Mass}_{measured} \text{ x } \frac{\mathrm{RU}_{nominal}}{\mathrm{RU}_{exact}} \tag{1.20}$$

$$\mathrm{KMD} = (\mathrm{Mass}_{nominal} - \mathrm{KM}) \text{ x } 1000 \tag{1.21}$$

Equation 1.20 determines an ions Kendrick mass (KM) by using the Mass of the measured ion and normalising it to the repeating unit (RU), causing each mass belonging to the same homologous series (i.e. additions of the RU) to have the same defect (decimal). This KM is then used to calculate the KMD, as per Equation 1.21, which is plotted against the nominal mass of the ion as shown in Figure 1.16. As highlighted by the horizontal dashed lines, this provides easy identification of masses belonging to the same homologous series. Kendrick mass analysis has been used often as a means of compound assignment, although this is rarely done now due to the advancement of automated assignment tools.[100,101]

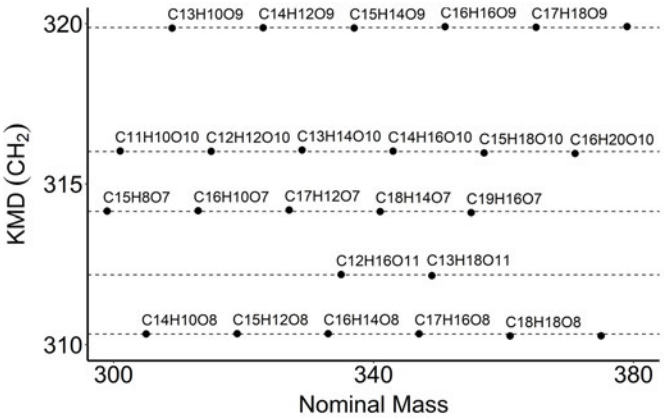


Figure 1.16: An example of a Kendrick mass defect plot applied to a NOM standard, here the repeating unit used is the CH_2 unit.

The data output from FT-ICR of complex mixtures is inherently complex itself, this makes interpretation and visualisation of the information extremely difficult. There are several methods that are often employed to achieve this and will be described here.

A method often used for the visualisation of bulk samples is known as a Van Krevelen diagram, originally published in 1950 as a method of visualising the average molecular formula of 3 different coal samples by plotting the hydrogen-carbon ratio (H/C) *vs* the oxygen-carbon ratio (O/C).[102] The Van Krevelen diagram was first employed for the analysis of NOM data in 2003 and has gone on to become one of the most popular forms of complex mixture MS data visualisation, with a lot of groups implementing classification systems dependant on where molecular formulae are positioned on this plot. An example Van Krevelen diagram for a NOM sample is given in Figure 1.17, with an example of one of these classification systems, highlighting that NOM consists of many lignin like structures.[103]

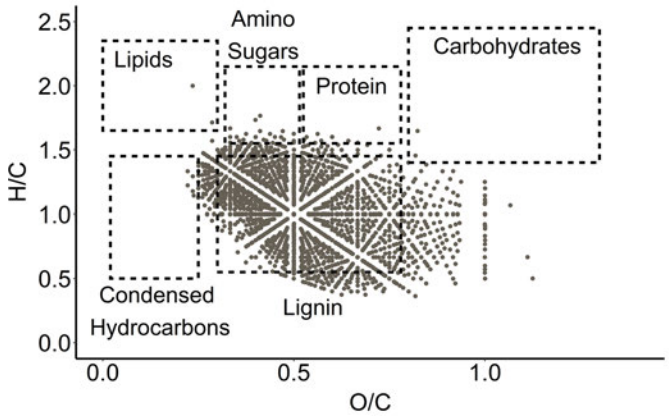


Figure 1.17: An example Van Krevelen diagram of a NOM sample, labelled boxes are an example of a classification system commonly applied to these plots.

It should be noted, that these classification systems are not absolute and it is possible for a compound belonging to one class to be positioned within the assignment area of another class.

Having assigned a molecular formula, there are several approaches one could take to inspect this data. One common application is to plot heteroatomic classes, such as an oxygen series plot, like that shown in Figure 1.18b. Another method is to calculate and plot the Double Bond Equivalence (DBE, Equation 1.22) of each molecular formula; this provides information on the degree of unsaturation in the molecule.

$$\mathrm{DBE} = \mathrm{C} + 1 - \frac{\mathrm{H}}{2} - \frac{\mathrm{X}}{2} + \frac{\mathrm{N}}{2} \tag{1.22}$$

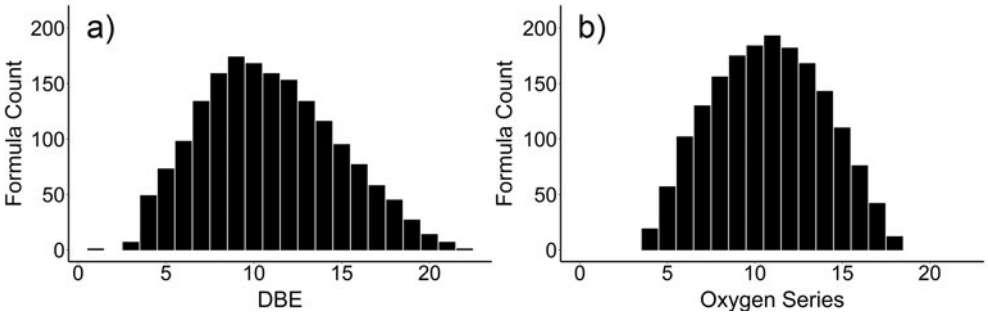


Figure 1.18: Example a) DBE and b) Oxygen Series plots for a NOM standard sample.

Many of the formulae assigned to complex mixtures analysed by negative mode ESI MS are abundant in oxygen, which is not a factor in the DBE equation, this has led some groups to produce classifications that are better suited to the analyses of NOM. One such parameter is known as the aromaticity index (AI), shown in Equation 1.23. This equation was later modified (AI_{mod})

based on the fact that most oxygen in NOM is present in carboxylic acid moieties, where only half of the oxygen contributes to conjugated systems, while the other half is σ-bound. This modification is shown in Equation 1.24.

$$\mathrm{AI} = \frac{1 + \mathrm{C} - \mathrm{O} - \mathrm{S} - \frac{1}{2}(\mathrm{N} + \mathrm{P} + \mathrm{H})}{\mathrm{C} - \mathrm{O} - \mathrm{S} - \mathrm{N} - \mathrm{P}} \tag{1.23}$$

$$\mathrm{AI}_{mod} = \frac{1 + \mathrm{C} - \frac{1}{2}\mathrm{O} - \mathrm{S} - \frac{1}{2}(\mathrm{N} + \mathrm{P} + \mathrm{H})}{\mathrm{C} - \frac{1}{2}\mathrm{O} - \mathrm{S} - \mathrm{N} - \mathrm{P}} \tag{1.24}$$

Both of these AI equations are aimed at assigning formulae into one of three groups, either non-aromatics, aromatics or condensed aromatic species, as shown in Figure 1.19.

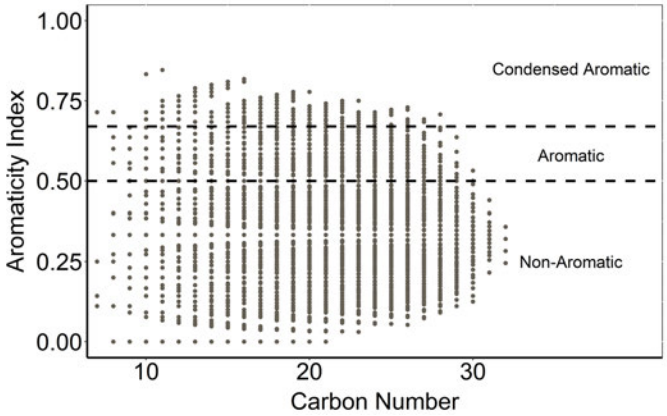


Figure 1.19: Example AI plot on a standard NOM sample, with dashed lines indicating the separation between the three classes, non-aromatic (AI $<$ 0.5), aromatic (0.5 $<$ AI $<$ 0.67) and condensed aromatic (AI $>$ 0.67).

A modification to the AI_{mod} classification is used throughout this thesis and was first published by Kellerman et al[104] to analyse NOM. The classification separates formulae into 4 distinct categories, Polycyclic aromatics ($\mathrm{AI}_{mod} >$ 0.66), plant-derived polyphenolics (0.5 $< \mathrm{AI}_{mod} \leq$ 0.66), highly unsaturated and phenolic species ($\mathrm{AI}_{mod} \leq$ 0.5 and H/C $<$ 1.5) or aliphatic species (AI_{mod} $\leq$ 0.5 and H/C $\geq$ 1.5). This system is arguably an improvement as it not only relies on the AI_{mod} but also takes elemental ratios into account through the use of the H/C ratio. An example of this classification is shown in Figure 1.20.[104]

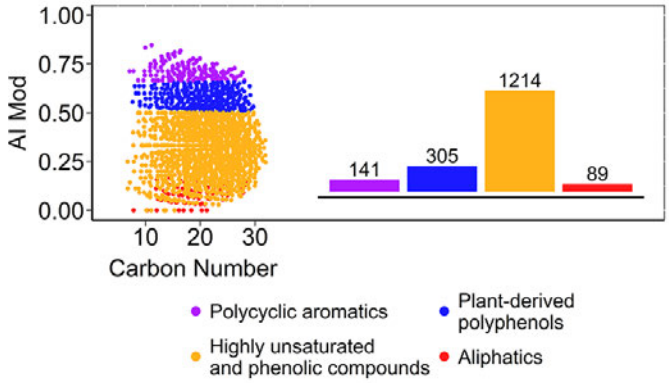


Figure 1.20: A Kellerman AI_{mod} plot highlighting the molecular characteristics of the molecular formulae present in SRFA.

AIMS

DOM has the potential to negatively impact almost every stage of potable water production. Its effective removal can therefore make water treatment works more efficient in producing higher quality drinking water. However, due to the complex nature of DOM, its fate during key treatment stages such as coagulation and ion-exchange, is not well understood. Increasing our understanding of the composition of DOM and its transformations throughout the potable water production process will provide valuable insights for its effective removal. Towards this end ultra-high resolution analytical techniques of Fourier transform ion-cyclotron resonance mass spectrometry (FT-ICR-MS) and nuclear magnetic resonance (NMR) were used as the main experimental methods.

The aims of this project are as follows:

- Develop and optimise FT-ICR-MS and NMR methods for the analysis of DOM.
- Investigate a pilot plant utilising Suspended Ion - Exchange (SIX) and In-Line Coagulation-Ceramic Membrane Filtration (ILCA-CMF) and characterise their DOM removal characteristics.
- Investigate which technique, Granular Activated Carbon (GAC) or Ion - Exchange (IEX), would be better suited as a post coagulation treatment.
- Characterise the composition of DOM as it undergoes photocatalytic degradation by TiO_2 and a hybrid of bismuth titanate ($Bi_4Ti_3O_{12}$) and TiO_2 in the presence of UV-vis irradiation.
- Develop ^{19}F based NMR experiments for extracting NMR parameters required for structure determination of fluorinated compounds contained in complex mixtures.
- Utilise ^{19}F NMR to investigate the disinfection by-products produced by chloramination of water containing DOM.

Part II

PILOT PLANT STUDIES

SUSPENDED ION EXCHANGE (SIX) PILOT PLANT

3.1 DECLARATION

Most of the work in this chapter has been previously published in the following paper:

Smith, Alan J. R. and Moore, Graeme and Semiao, Andrea J. C. and Uhrín, Dušan. Molecular level characterisation of ion-exchange water treatment coupled to ceramic membrane filtration. *Environmental Science: Water Research & Technology*, 2020.

http://dx.doi.org/10.1039/C9EW01042D

Sample collection and data acquisition were performed by myself, however, Scottish Water operated the pilot plant.

3.2 OVERVIEW AND PILOT PLANT SETUP

Dissolved Organic Matter (DOM) can be a source of numerous issues within a water treatment process. It can act as a source of energy for microorganisms resulting in regrowth throughout the distribution network; not only is this a potential biohazard but it also increases the turbidity of the water. This regrowth results in the formation of a biofilm, something usually present in all distribution networks, however if hydraulic conditions in the distribution network change, there is a potential to release this biofilm resulting in a degradation of water quality at the consumer output.[105]

The presence and persistence of DOM throughout the treatment process does have another significant drawback in that many of the most commonly used disinfectants have the ability to react with DOM to produce disinfection by-products (DBP's). Due to the shear complexity of DOM as a precursor the resulting by-products are often difficult to detect (beyond the most basic) and as a result, many remain unknown along with the potential effects to human health.

Many tactics are employed by water suppliers trying to prevent the breach of DBP regulations, such as optimising the disinfectant dose, changing the method of disinfection entirely or introducing active removal of by-products post treatment. One of the most attractive strategies involves the reduction of DOM prior to the disinfection process, so the DBP's are not produced to begin with.

Several processes have been researched for this purpose, such as filtration via membranes prior to treatment, the use of granular activated carbon (GAC) filters and ion-exchange processes (IEX). Ion-exchange in particular has received significant attention and for good reason, in many cases it has been shown to result in high DOM (measured as DOC) removal, with several groups indicating that this process targets fractions of DOM that coagulation cannot remove.

To that end, this chapter involves the molecular level characterisation of DOM using ultra-high resolution mass spectrometry and high-field NMR to investigate the effect of combining a novel Suspended Ion-Exchange (SIX) process with In-Line Coagulation (ILCA), followed by a final filtration step using a ceramic membrane (CeraMac) at a pilot plant scale. The plant was located in the north east of Scotland, produced by PWN Technologies (Netherlands) and was capable of producing 150 m^3/day. The plant schematic is shown in Figure 3.1.

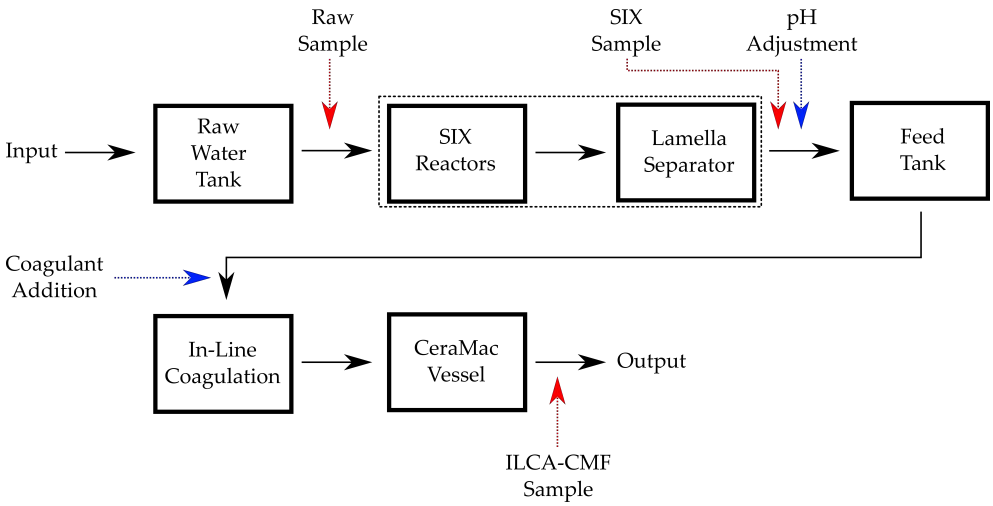


Figure 3.1: Schematic highlighting main processes occurring throughout the SIX based pilot plant. Red arrows represent sampling locations and blue arrows indicate the location of pH adjustment or coagulant addition. Dashed box area is expanded in Figure 3.2.

The inlet water is first exposed to the IEX resin within the SIX reactors, where the IEX resin rather than being a uniform bed of polymeric beads is held in suspension within the water by bubbling air from the bottom. This allows

more of resins surface to be in contact with the water increasing its efficiency. The resin is then retained via a lamella separator, where it is either recirculated or regenerated using a 30 g/L NaCl solution (every 30 minutes). The resin used was an anionic resin in chloride form (Lewatit S5128 Germany) and during this experiment, had been in circulation for 5-7 months.

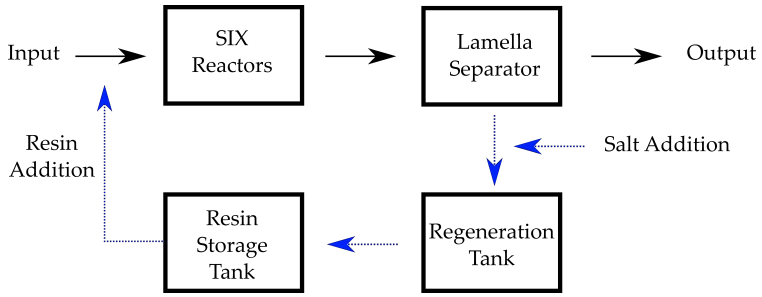


Figure 3.2: Expanded schematic of the SIX element of treatment. Blue arrows highlight the movement of the resin throughout the process.

After having passed through the SIX, the water then undergoes in-line coagulation (ILCA) using polyaluminium chloride (Water Treatment Solutions, UK) at a pH of 6.4. The pH corrections were made using NaOH or HCl, injected with the coagulant and mixed with a static mixer before flocculating for approximately 4 minutes. Optimal coagulant dose and pH adjustments were established by jar testing on site and through zeta potential analysis using a zetasizer (by Scottish Water).

Finally, the coagulated water is passed through a ceramic membrane element (Metawater, Japan) which is a vertically mounted 25 m^2 element with a nominal pore size of 0.1 μm. The element was operating via dead-end filtration and was backflushed every 10 minutes in order to prevent biofouling.

The adjoining treatment works also used polyaluminium chloride as a coagulant, flocculation mixers for in-line flocculation and ultrafiltration (UF) using Memcor hydrophilic membranes (polymeric ultrafiltration membranes) with a pore size of 0.045 μm, a schematic can be found in Figure 3.3.

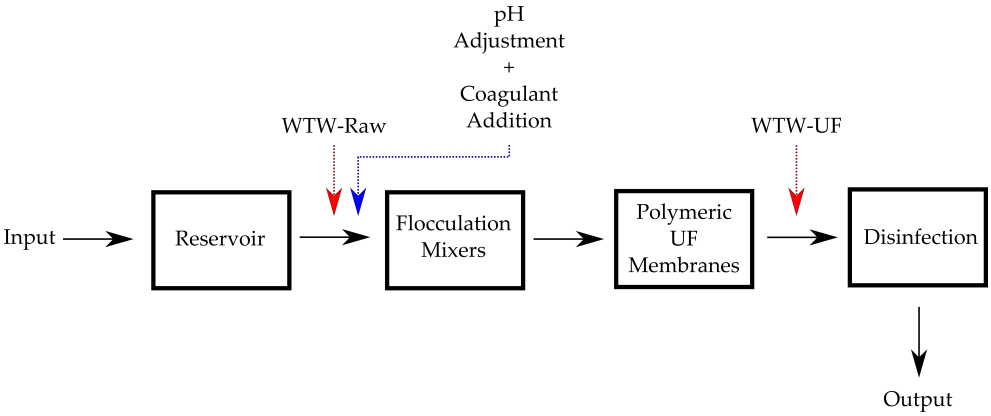


Figure 3.3: Schematic highlighting main processes occurring throughout the WTW adjoining the SIX based pilot plant. Red arrows indicate sampling locations and blue arrows indicate location of pH adjustment or coagulant addition.

3.3 EXPERIMENTAL METHODS

Due to the layout of the pilot plant at the time of sampling, it was only possible to sample the raw inlet water prior to SIX, after SIX and the final output water after ceramic membrane filtration. Samples were collected in acid washed amber glass bottles from the pilot plant on two dates, the 12th of June 2017 and the 17th of July 2017, and once from the adjoining water treatment works, also on the 17th of July. Sample volumes are as follows -

i) 2.5 L for raw inlet water, the sample is termed Raw if from the pilot plant or WTW - Raw if from the water treatment works.

ii) 5 L for the water prior to SIX treatment, this sample is herein termed SIX.

iii) 10 L for any final output water, termed either as ILCA - CMF if from the pilot plant or WTW - UF if from the water treatment works.

Water samples were concentrated, desalted and freeze-dried so that sufficient amount of DOM was obtained and there would be minimal salt interference when interpreting complex mass spectral data. Samples were filtered through Whatman ME25 mixed cellulose ester filters (0.45 μm) on the day they were collected. Solutions were then concentrated using a custom built cross-flow reverse osmosis (RO) system, utilising BW30 membranes (DOW Filmtec), with a cross flow rate of approximately 1 litre per minute and a pressure between 13 - 15 bar. Solutions were concentrated approximately 20x before being passed to a commercially available electrodialysis system (PCCell) to remove monovalent

and divalent salts. The system used was a two chambered cell design, containing 10 cell pairs with a total membrane area of 0.25 m^2, each cell pair consisted of both a cation and anion exchange membrane. The process was stopped when the initial sample conductivity (between 250 - 350 μS) had dropped to less than 50 μS. Samples were then lyophilised and stored in glass vials for further analysis.

A lab experiment was conducted wherein the resin in use at the pilot plant was exposed to a raw DOM solution throughout multiple use/regeneration cycles. The IEX resin was exposed to a solution of DOM at a concentration of 10 mg/L and stirred for 30 minutes, before being stirred in a solution of NaCl at 30 g/L (the same as used in the pilot plant itself), before being rinsed in deionised water for a further 15 minutes. This was repeated six times.

For (-) ESI FT-ICR-MS analysis samples were dissolved to a concentration of 0.1 mg ml^{-1} using a 50:50 mix of MS grade MeOH and D_2O obtained from Sigma Aldrich Co and introduced to the spectrometer using a syringe pump at 200 μL h^{-1}. The ion accumulation time and time of flight were set to 0.25 sec and 0.6 ms respectively. Spectra were acquired at 4 MW resulting in an FID length of 1.666 secs, 600 transients were accumulated.

For the (-) LDI samples, IEX beads were 'attached' to a recessed MALDI plate using double sided copper tape. 100 laser shots at 25% power were fired for each of the 200 transients acquired. Samples were acquired as a Ser file, allowing the removal of transients without significant ions. Ser files were opened using the Bruker software FTMS-Processing, where transients were inspected for data quality. Ser files were then opened in python using the python package NMRGlue, where the selected transients were then summed to create an FID, which when returned to FTMS-Processing can be transformed to provide a mass spectrum. The python script for separating the ser files is given in appendix Figure A.1.

Masses between 200 and 600 m/z were peak picked and utilised further throughout the course of this investigation as this was the region which provided the highest S/N ratio across the mass range acquired (147 - 3000 m/z). Formulae were assigned using in-house scripts written by Dr Will Kew, a former student in the DU group, with constraints of $^{12}C_{0-66}$ $^{1}H_{0-126}$ and $^{16}O_{0-27}$.[106] Only CHO containing species were investigated as nitrogenous species are typically not well represented in negative mode ESI spectra[107] and the resolution required to confidently assign these species is greater, leading to the need for increased FID lengths and thus longer overall experiment times.

For NMR analysis, lyophilised DOM was dissolved to a concentration of 1 mg 600 μl^{-1} using 99.9% D_2O acquired from Sigma Aldrich Co. Spectra were

acquired on an Avance III HD 600 MHz Bruker spectrometer equipped with a 5 mm TCI cryoprobe. Spectra were acquired using the NOESY based water suppression technique with the relaxation times and acquisition times set to 8.3 and 1.3 seconds respectively. 128 scans were acquired per spectrum before being zero filled to 128k points and having an exponential line broadening of 5 Hz applied prior to Fourier transformation.

ATR-FTIR spectra were acquired on a Perkin Elmer Spectrum Two spectrometer between 450 and 4000 cm^{-1} with a resolution of 0.5 cm^{-1}. Spectra were acquired in triplicate and a standard normal variate correction was applied to each individual spectrum. Relative signal intensities were calculated individually for each spectrum as T_i / $\Sigma(T_i)$ where T_i is the transmittance at an individual wave number. PCA on these spectra was performed using SIMCA version 14.1 (Umetrics), using pareto scaling on wavenumbers between 800 - 2000 cm^{-1}.

3.4 RESULTS AND DISCUSSION

3.4.1 *TOC Analysis*

TOC measurements were taken at the time and point of sampling, the results of this are presented in Table 3.1. Despite the much larger TOC levels in June, the combined SIX and ILCA-CMF treatment was able to lower the TOC to a similar value (8% and 14%) to that seen in July. The ILCA-CMF was approximately twice as efficient in DOM removal as the SIX treatment.

Scottish Water provided the total organic carbon (TOC) values presented in Table 3.1.

Table 3.1: TOC values of the collected samples.

Sample	TOC (mg C/L)	
	June	July
Raw	8.6 ± 0.23	3.5 ± 0.22
SIX	5.4 ± 0.26	1.8 ± 0.07
ILCA - CMF	0.7 ± 0.18	0.5 ± 0.8
WTW - Raw	-	3.7 ± 0.11
WTW - UF	-	0.9 ± 0.16

[a] standard deviation is based on five injections.

The DOM retention by reverse osmosis was assessed by measuring the TOC after the RO concentration step by diluting a portion of the sample to the appropriate original volume. The yields found were 92±3%. This method of concentration generally achieves much higher levels of organic matter retention than other commonly used methods such as SPE.[39,108]

3.4.2 *(-) ESI FT-ICR-MS Analysis*

During this experiment samples were analysed using ESI (-) mode ionisation as this is the easiest and most commonly used ionisation technique for the analysis of complex organic matter samples, a typical spectrum is shown in Figure 3.4.

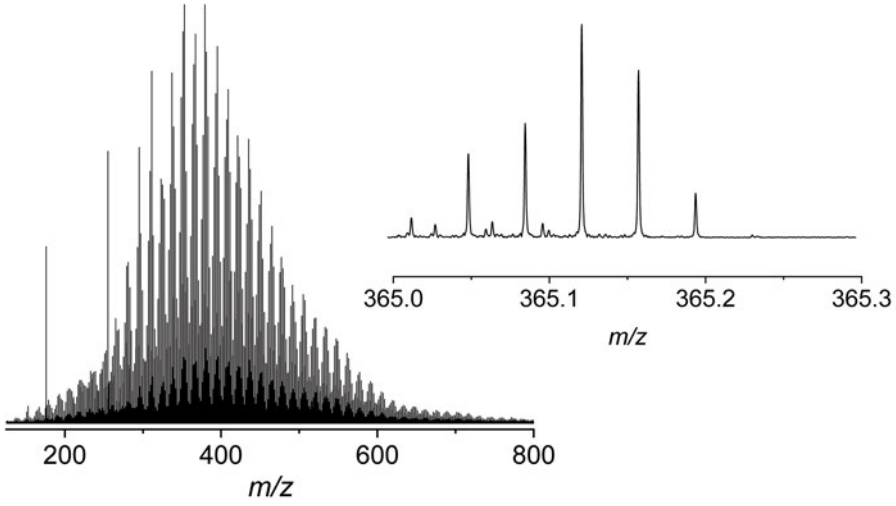


Figure 3.4: Representative negative mode ESI spectrum of raw water DOM with a zoomed region at 365 m/z.

The assignment statistics are shown in Table 3.2. The average assignment rate was 81.5 (±5)%.

From Table 3.2, it can be seen that the raw and SIX treated water samples have very similar assignment efficiency despite the drastic difference in TOC values presented in Table 3.1. This makes the considerable drop in formula assignment after the ILCA-CMF treatment all the more interesting, suggesting that the treatment process is drastically altering the organic matter composition.

In order to investigate differences (or lack thereof) between samples, intersections of assigned formulae were examined using UpSet plots.[109] These plots allow the graphical visualisation of the number of formulae common to differing subsets of samples within a given sample set. The portion of the figure labelled "Set Size" indicates the total number of formula assigned in that sample. Dots right of the sample name indicate the presence of formulae in that sample, i.e. intersects,

Table 3.2: Summary of FT-ICR-MS Analysis

Sample	Total Peaks Picked		Monoisotopic		Isotopic		Assigned (%)	
	June	July	June	July	June	July	June	July
Raw	4956	4845	2518	2427	1364	1301	78	77
SIX	4217	4440	2568	2543	1104	1072	87	81
ILCA-CMF	3482	2885	1981	1656	920	631	83	79
WTW - Raw	-	4083	-	2393	-	1091	-	85
WTW - UF	-	4090	-	2266	-	1079	-	82

[a] Isotopic formulae are ^{13}C equivalents of the monoisotopic assignments.

of which there are a maximum of 2^n - 1, where n is equal to the number of samples. As an example, looking at Raw - July in Figure 3.5a, that sample has the sum of the dots on the right, i.e. $2004 + 133 + 79$.

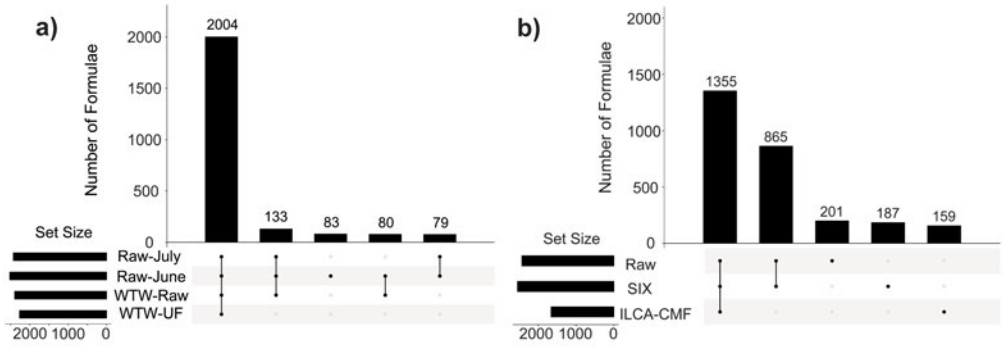


Figure 3.5: **a)** UpSet plot highlighting the similarities between the Raw water and WTW - UF sample and **b)** UpSet plot showing the differences between the raw/SIX and the ILCA-CMF samples taken in July, an equivalent plot for June can be found in Figure A.2.

Due to the similarity in assigned formulae numbers given in Table 3.2, the raw pilot plant water samples were initially compared with the raw and ultrafiltrated (UF) waters from the WTW, shown in Figure 3.5a. As can be seen there is very little difference between the three raw water samples, this is surprising as it shows that although the overall organic levels (TOC) are drastically different, the composition at the molecular level (accessible by (-) ESI at least) is very similar. This figure also shows that the coagulation and ultrafiltration treatment in the WTW did not alter the molecular composition to any significant extent.

This wasn't expected as the TOC after treatment has dropped considerably and these techniques are known to result in the preferential removal of species with high O/C ratios.[110,111] An implication is, that although the TOC level is decreasing, it appears to be removed non selectively, or more likely, those species which are preferentially removed are not the species observable using (-) ESI FT-ICR-MS, hence why NMR, a less selective analytical technique is used later in this chapter.

Investigating exclusively the July pilot plant samples (3.5b) highlighted the extent of dissimilarity between the ILCA-CMF samples and the raw/SIX treated waters, emphasising that the ILCA-CMF combined treatment step is resulting in considerable changes to the molecular composition of the DOM. An identical plot for the June samples is shown in the appendix Figure A.2. This plot also reveals a large similarity between the raw water and the SIX treated water samples (93.2% similarity in July, 91.5% in June). This suggests that either the SIX treatment is removing species non-selectively or that the fraction removed is not observable using this ionisation technique. Experiments performed on a similar pilot plant setup using liquid chromatography - organic carbon detection (LC-OCD), have suggested that the SIX process was selectively removing lower molecular weight species, which does not appear to be the case in this investigation.[112]

To highlight just how different the ILCA-CMF treated sample is, in comparison to the raw and SIX treated waters, Figure 3.6 shows a zoomed region of the three spectra (365.0 - 365.3 m/z). It can be seen that the species furthest left in the distribution, those which are often attributed to more aromatic species, are missing or substantially reduced in intensity, this holds true across all m/z values indicates that the ILCA-CMF treatment is selectively removing species with a low H/C ratio. This figure also provides additional evidence to the apparent lack of selectivity in DOM removal by the SIX treatment, i.e. the two top spectra are practically identical.

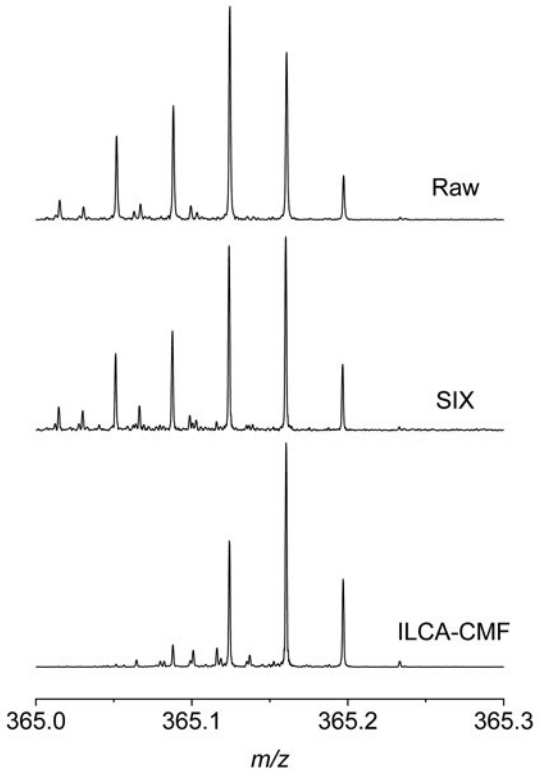


Figure 3.6: Overlay of three mass spectra, zoomed to the 365.0 - 365.3 m/z region, highlighting the difference in molecular composition after ILCA-CMF treatment.

In order to visualise this difference in terms of compound classes, Van Krevelen diagrams were also produced (Figure 3.7). These diagrams show individual molecular formulae as points on a coordinate system, hydrogen/carbon ratio (H/C) *vs* oxygen/carbon ratio (O/C), with regions in this coordinate system being assigned to different compound classes. This method is not unambiguous, but it does allow for a quick visualisation of the molecular species within a sample.

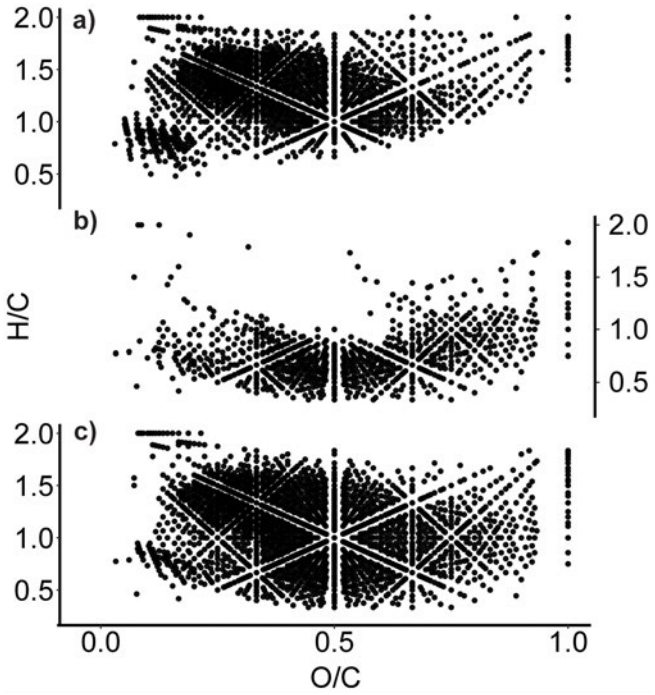


Figure 3.7: Van Krevelen diagrams of **a)** the June pilot plant sample after the ILCA-CMF treatment; **b)** formulae that are no longer present after the ILCA-CMF treatment and **c)** the June SIX sample. An equivalent plot for the samples taken in July is shown in Appendix Figure A.3

Figure 3.7 shows a comparison of Van Krevelen diagrams focusing on the SIX and ILCA-CMF treated samples. As can be seen in 3.7b, the majority of the removed compounds have a H/C ratio of between 0.5 and 1, typical for highly aromatic species. A large portion of these species also have relatively high oxygen content. This could indicate that the species are polyphenolic carboxylic acids, a class of compounds which have been shown to produce higher levels of disinfection by-products than the less aromatic species.[113]

Another metric used to characterise differences between samples is known as the modified Aromaticity Index (AI_{mod}), a classification which categorises compounds as non-aromatic, aromatic and condensed aromatic based on their molecular formula. Figure 3.8, shows the AI_{mod} plot produced for the July samples, each sample has been normalised to 100%. As can be seen, this metric shows very little difference between the raw water and the SIX treated water samples as indicated by previous analysis.

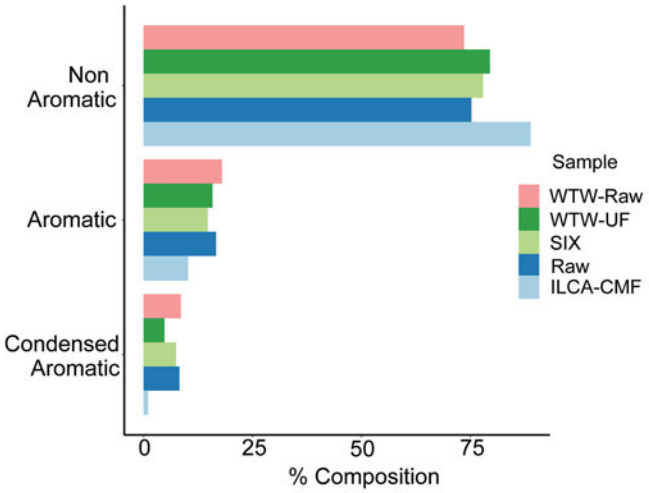


Figure 3.8: AI_{mod} plot of the SIX based pilot plant, samples shown were those taken in July, an equivalent plot for those taken in June can be found in Appendix Figure A.4.

The AI_{mod} also highlights the extent of the removal of aromatic species via the ILCA-CMF treatment. In comparing the numbers of formulae assigned the SIX treated sample and the ILCA-CMF treated sample, there is a 91% loss of species which were designated as condensed aromatics and a 54.8% loss of aromatic species. Interestingly, despite the degree of similarity between the WTW-Raw and WTW-UF sample highlighted in Figure 3.5a, the AI_{mod} plot revealed a significant reduction in the number of species designated as condensed aromatic, 47.1% less than those present in the WTW-Raw sample. This wasn't easily noticeable using the UpSet plots as it only amounts to a difference of 96 formula.

As mentioned earlier, the Van Krevelen plot highlighted that a large proportion of oxygen containing species had been removed by the ILCA-CMF treatment. In order to investigate these further, oxygen series were plotted to visualise the oxygen distribution across the treatment stages. As can be seen in Figure 3.9, the proportion of assignments above O_6 started to decline after the ILCA-CMF stage, whilst those in the raw and SIX treated waters continued to increase. This decline further accelerated at O_{11} and the depletion of these higher oxygen containing species became more prominent.

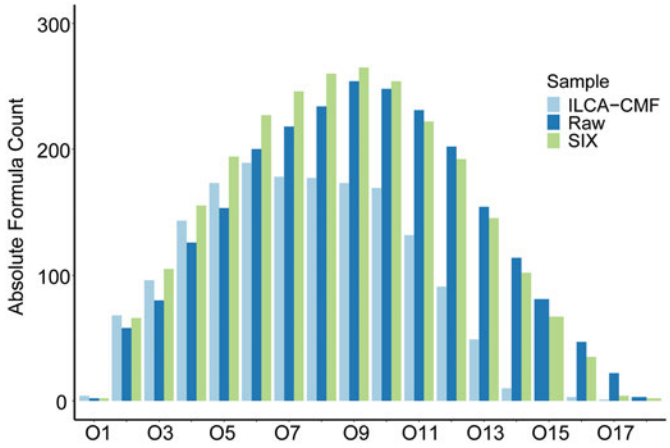


Figure 3.9: Oxygen class plot of the July pilot plant samples. Equivalent plot for June samples is shown in Figure A.5.

To summarise, the FT-ICR-MS data shows that despite the SIX treatment causing a drastic reduction in the TOC levels, it was non-selective in terms of the species removed, at least considering small molecules which are ionised via (-) ESI mode. The MS analysis has also shown that the ILCA-CMF treatment is selectively removing the more aromatic/condensed aromatic species with high oxygen content.

3.4.3 ^{1}H *NMR Analysis*

Despite NMR spectroscopy being an extremely powerful technique for the structural determination of pure compounds or small molecules within the most simple of mixtures, it does fall short when analysing more complicated systems, especially when these mixtures are composed of several hundreds or thousands of unknowns. NMR lacks both the sensitivity and the resolution to be able to discern structural characteristics within these complex organic matter samples, nevertheless, it is non-selective and the dispersion of the ^{1}H chemical shifts and the quantitative nature of NMR means that it can be used to estimate the compound class distributions at the various stages of sampling.

^{1}H NMR spectra of DOM samples (Figure 3.10) show just how complex these samples are; ultimately the apparent broadening of the signals is caused by many molecular species having chemical shifts slightly offset from each other.

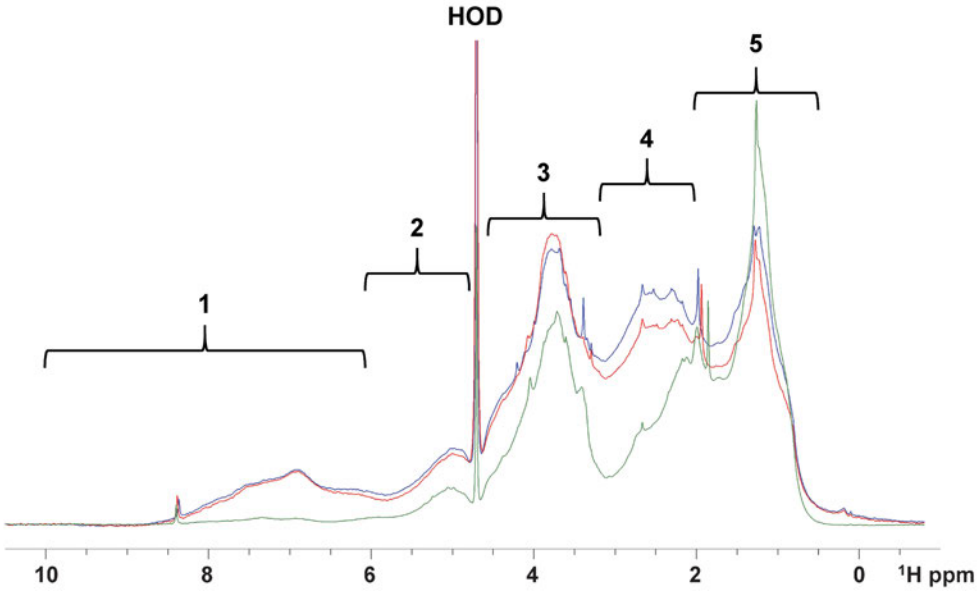


Figure 3.10: Superimposed NMR spectra of pilot plant June samples for raw (blue), SIX (red) and ILCA-CMF (green). Numbered sections indicate the chemical shift regions as **1** - aromatic species (6 - 10 ppm), **2** - unsaturated region (4.8 - 6 ppm), **3** - carbohydrate region (3.1 - 4.6 ppm), **4** - Carboxyl Rich Alicylic Molecules (CRAM, 2 - 3.1 ppm) and **5** - aliphatic region (0.5 - 2 ppm). HOD indicates the signal from residual water present in the solvent.

In order to interpret these ^{1}H NMR spectra more effectively, a previously described classification[114] was adapted and used to produce integrals of five spectral regions shown in Figure 3.10. Their visual inspection clearly showed that the ILCA-CMF spectrum is characterised as having an increased intensity in the 1-2 ppm region, decreased intensity in the 2-3 ppm range and an almost complete disappearance of any signals in the aromatic region of the spectrum between 6-9 ppm. The spectra of the raw and SIX treated samples are very similar as was expected from the previous MS analysis, although a slight drop in the Carboxyl-Rich Alicylic Molecules (CRAM, 2 - 3.1 ppm) region is evident after having performed SIX treatment.

Integral intensities of individual spectral regions are summarised in Figure 3.11. It can be seen that the SIX treatment did not alter the ratios of the individual proton regions substantially. After the ILCA-CMF treatment however, the relative amount of aliphatic compounds has increased, while the amount of carbohydrates has reduced. Aromatic species after the ILCA-CMF treatment are significantly depleted, they represent only 1% of the total signal intensity in the ILCA-CMF samples compared to 9% in the untreated, raw water. The ^{1}H NMR spectra of the WTW samples also corroborate the result from the mass spectrometry. Here a reduction from 9% to 5% was observed for the aromatic

region, i.e. much less of a reduction when compared to final output water from the pilot plant.

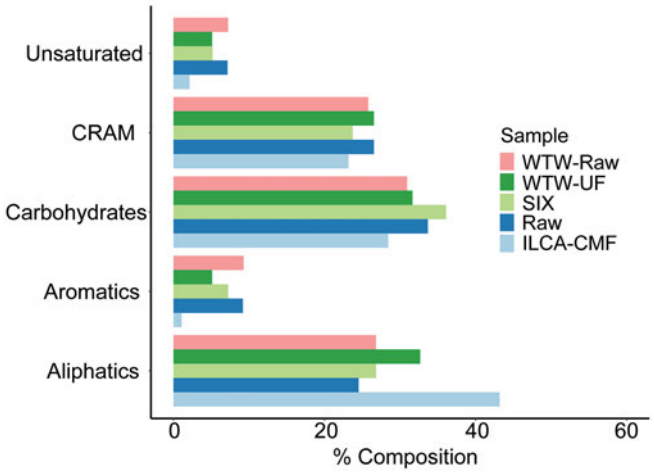


Figure 3.11: NMR integration results for the July pilot plant samples and the WTW samples. The equivalent plot for the samples taken in June is shown in appendix Figure A.6. The bar order is consistent with the order in the legend.

Interestingly, the NMR results follow the same compound distribution and changes throughout the treatment stages as the mass spectrometry, despite much higher detection limits. This indicates that there is a genuine and significant reduction in aromatic species after the ILCA-CMF treatment, and it's not just an artefact of the targeted analysis produced by the FT-ICR-MS.

3.4.4 *FT-IR Analysis and PCA*

Having seen evidence that the majority of the compounds which were removed by the ILCA-CMF treatment were aromatic in nature, ATR-FTIR was used in order to determine if this economical and much more accessible technique could be used to obtain the same result. Twenty four ATR FT-IR spectra of freeze dried samples (triplicates of the eight samples shown in Table 3.1) were acquired. An overlay of three representative spectra is shown in Figure 3.12. While the spectra for the raw and SIX samples are very similar, there are clear differences shown in the samples after ILCA-CMF treatment. The largest reduction in intensity occurred between 1800 and 1250 cm^{-1}, the region corresponding to the stretching vibrations of carbonyls, alkenes, arenes and also the bending vibrations of CH_2, CH_3 groups, OH and COH bending. This is consistent with the loss of more aromatic species and a reduction in those species possessing larger numbers of OH groups. There is also an increased intensity around 1060

cm^{-1} corresponds to C-O stretching and can be explained by the increased percentage of carbohydrates in this sample.

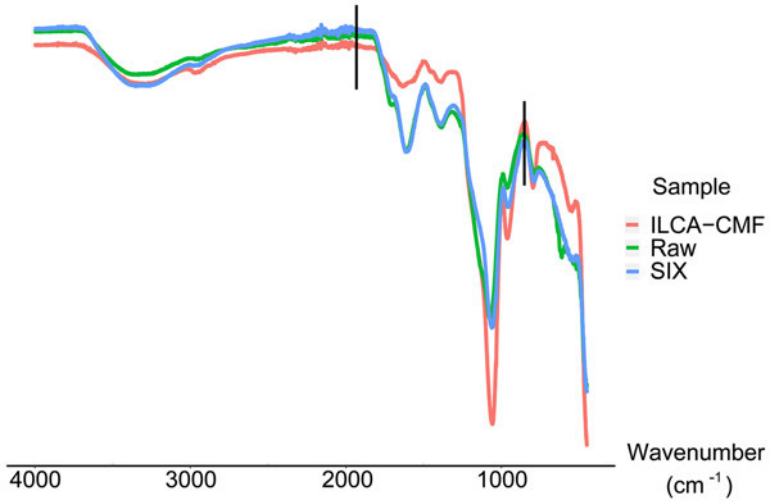


Figure 3.12: ATR-FT-IR spectra of July pilot plant samples. The region indicated by the two vertical black lines was used for PCA analysis.

To identify the more subtle differences between spectra, principle component analysis (PCA) of the relative signal intensities, calculated individually for each spectrum as $\frac{T_i}{\Sigma\, T_i}$ where T_i is the transmittance at individual wave numbers between 800 and 2000 cm^{-1}. The region lower than 800 cm^{-1} was removed due to the presence of instrumental noise and the region above 2000 cm^{-1} was excluded as it was featureless, apart from a broad band centred at 3400 cm^{-1} (H-bonded OH stretch), which could be affected by varying degrees of moisture within the samples.

The PCA analysis yielded an initial principal component which explained 81% of the data (PC1), with only 6% being explained by the second component (PC2). The corresponding scores plot is shown in Figure 3.13a. The replicate samples all group closely along PC1 and PC2 indicating that this method has a good reproducibility.

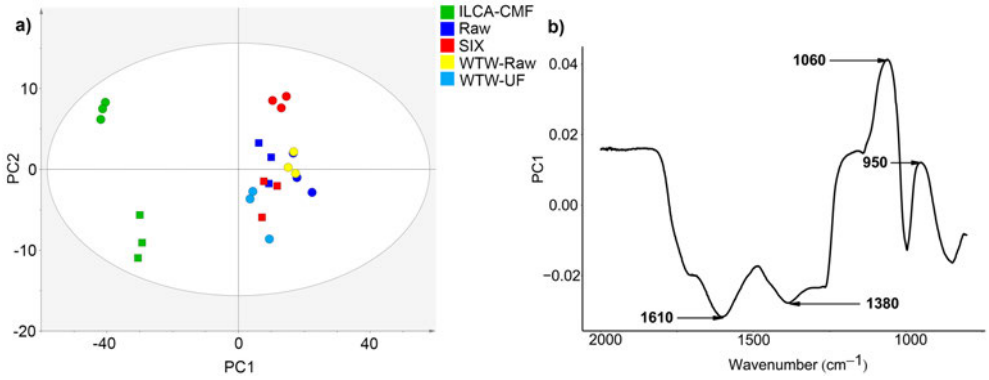


Figure 3.13: **a)** PCA scores plot of ATR-FTIR data for triplicates of the samples listed in Table 3.1. The Hotelling's T2 ellipse represents 95% confidence interval. The June and July samples are represented by squares and circles respectively; **b)** loadings plot for PC1 vs wavenumber.

Both the July and the June pilot plant ILCA-CMF treated samples are clearly separated from the rest of the sample set along PC1. As the intermonth variation along PC2 often exceeds the variation between samples receiving different treatments, it becomes difficult to discern whether the information in this component can contribute any more information, this is something that would need to be investigated on a larger dataset.

The loadings plot for PC1 (Figure 3.13b) indicates which variables are contributing to the definition of PC1 (negative values - i.e. those which decreased in ILCA-CMF samples) are the wavenumbers 1380 cm^{-1}, corresponding to phenolic, CH_2 and CH_3 deformations,[115] 1610 cm^{-1} which corresponds to olefinic and aromatic C=C[116] and 1710 cm^{-1} which can be attributed to the C=O stretch.[117] All of this indicates a loss of aromatic, carbonyl containing and polyphenolic species, the same conclusions as drawn from the two previous techniques, NMR and FT-ICR-MS. The positive loadings (i.e. those that increased following ILCA-CMF treatment) corresponds to 1060 cm^{-1}, C-O stretching,[118] which can be attributed to the increased relative percentage of carbohydrates in these samples, in comparison to other chemical species (note that there is still a drop in carbohydrates compared to the raw and SIX treated samples). The IR spectra are therefore in agreement with the ^1H spectra, where the resonances assigned to aromatic/olefinic groups and CH_2 groups were seen to be depleted.

As both the pilot plant and the WTW plant use the same coagulant and the WTW did not show any selectivity when processing the same inlet water, removal of aromatics/polyphenols in the pilot plant can therefore be attributed to the CMF step. A previous study,[112] used the same combination of processes, SIX and ILCA-CMF to assess their ability to remove disinfection by-product

precursors. They found via LC-OCD that SIX treatment preferentially removed low molecular weight species, while ILCA-CMF was more effective in removal of high molecular weight species. As discussed earlier these findings cannot be corroborated by FT-ICR-MS due to inability of this technique to observe simultaneously large and small compounds (in large numbers). Indeed, we saw little difference between the SIX treated and the raw water DOM via this technique. NMR spectroscopy would be able to observe proton signals from these larger molecular weight species, yet there is still little difference in the spectral profile. This previous study[112] did find however, that the potential to form disinfection by-products was greatly reduced after CMF; this could be explained by the removal of the aromatic species seen in this study.[119]

Similar to our observations, high-resolution MS, differential absorbance and fluorescence were able to detect changes in DOM composition, which could not be detected with commonly used DOC-normalized parameters, emphasising the usefulness of high-end analytical techniques in assessing the efficiency of new water treatment technologies.[120]

As demonstrated throughout this chapter, while FT-ICR MS provides the most comprehensive, yet only qualitative information concerning small and medium size molecules, ^{1}H NMR is a quantitative, low resolution technique, capable of unambiguously identifying the lack or presence of aromatic compounds. ATR-FTIR on the other hand is the most economical technique that is also sensitive to these types of compounds, although the overlap between the IR absorption bands can prevent unambiguous identification of structural fragments, a fact only worsened when analysing complex mixtures.

3.4.5 *IEX Lab Experiment - (-) LDI MS of SIX beads*

The previous results obtained from the SIX based pilot plant indicated that the DOM following SIX treatment was very similar to the untreated sample. The technique with the highest resolution is undoubtedly the FT-ICR-MS, it is however let down by the fact that the ionisation source, in this case (-)ESI, is a compound type dependant technique. Therefore, in order to further investigate the removal of compounds by the ion exchange resin, a second ionisation source, Laser Desorption Ionisation (LDI) was used. This technique is capable of ionising compounds adsorbed to the surface of the beads.

Samples from the IEX resin lab experiment were taken before every regeneration cycle and freeze dried, the IEX beads were 'attached' to a recessed MALDI plate using double sided copper tape as seen in Figure 3.14.

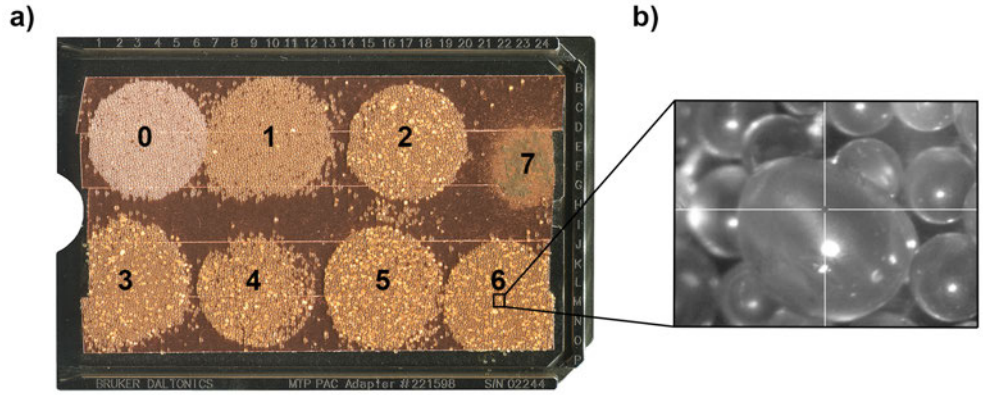


Figure 3.14: **a)** MALDI plate with immobilised OM loaded SIX resins (0-6), sample 7 is the raw DOM used throughout the adsorbtion experiment. **b)** View of the resin via the laser window on the FT-ICR-MS.

Sample 0 corresponds to the blank IEX bead, whilst samples 1 - 6 represent the number of use and regenerative cycles that bead has experienced, i.e. sample 6 has been exposed to the raw DOM sample 6 times and regenerated 5 times, sample 7 is the solid DOM sample used throughout the entire experiment. In order to combat some of the disadvantages associated with LDI, such as shot-to-shot reproducibility, a method was devised which would allow the user to select which transients to sum, an option currently unavailable in Brukers FTMS processing software. This method is outlined in the experimental details section of this chapter.

The number of formulae assigned in LDI - FT-ICR-MS spectra (Figure 3.15), indicates that samples 1 and 2 are underloaded and do not have a sufficient concentration of DOM for analysis, whereas sample loading appeared to level off after 4, 5 runs.

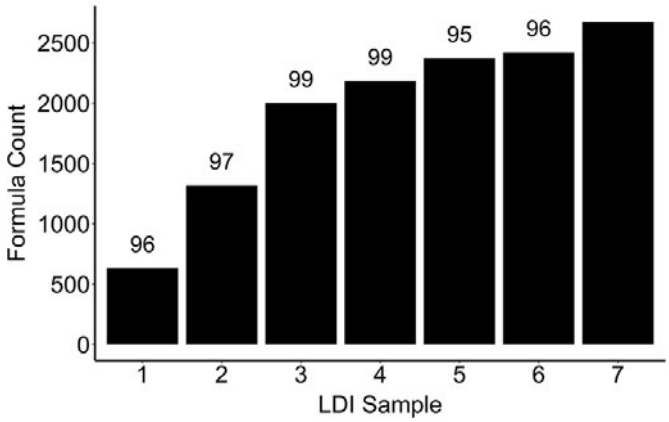


Figure 3.15: Bar plot highlighting the total number of assigned formulae in samples 1 - 7 described in Figure 3.14. Numbers situated above the bars indicate the percentage of formulae in each respective sample that are also in the raw DOM sample, number 7.

Figure 3.15 shows that only after 4+ runs, was there enough adsorbed material to produce a spectrum with a similar number of formulae as the raw DOM sample (sample 7). This is likely a result of a reduced concentration of DOM on the runs that were regenerated less times. There are 362 formulae present in the raw DOM sample vs sample 6, however these species have no obvious trend, i.e. no correlation with mass, DBE, oxygen series or AI_{mod}. This difference in formulae is likely to go down with further runs, as an increased concentration of irreversibly adsorbed material is deposited onto the bead. The high similarity of sample 6 and the raw DOM sample indicates again that the SIX process with this particular resin does not fractionate or preferentially remove a particular species type, i.e the DOM on the IEX beads appears to have the same composition as the raw DOM. This result is the same as observed with the (-) ESI method.

3.5 CONCLUSION

The application of (-) ESI FT-ICR-MS, NMR and ATR-FTIR provided independent evidence for selective removal of aromatic/phenolic species by the in-line coagulation and CeraMac filtration (ILCA-CMF), while the suspended ion exchange (SIX) treatment was found to be non-selective in the species removed as shown by these techniques and (-) LDI FT-ICR-MS. Near complete removal of these compound classes following a combined treatment by SIX and ILCA-CMF is significant, as these species are known to react readily with chlorine and are a proven precursor to trihalomethanes. As the SIX removes species in a non-targeted manner, it can be implemented prior to coagulation to reduce

the load on the coagulation and filtration system, it also has the potential to be used as a post filtration treatment, removing those species which the coagulation and filtration couldn't remove.

PILOT PLANT INVESTIGATING GAC AND IEX AS POST COAGULATION TREATMENTS

4.1 OVERVIEW AND PILOT PLANT SETUP

Unlike the previous pilot plant study (chapter 3), this pilot plant was designed to directly measure two processes against each other. Water from the existing treatment works is diverted before disinfection into a pilot plant, where it is filtered through both a Granulated Active Carbon (GAC) filter and an ion-exchange (IEX) resin bed in parallel. The aim of this pilot plant is to observe which of the two treatment processes is more advantageous as a post coagulation treatment.

The pilot plant is located at the site of an existing WTW in the central highland region of Scotland. The existing treatment works initially utilises micro straining to remove any particulate matter larger than 80 microns across before being coagulated using polyaluminium chloride. The water is then filtered using polyethersulphone ultrafiltration (UF) membranes (Norit Membrane Technology) with a pore size of approximately 0.02 μm. It is after this filtration stage that a portion of the water is diverted towards the pilot plant. After the ultrafiltration the water undergoes chemical conditioning, including pH adjustment using lime and phosphoric acid, before being disinfected using chlorination. The treatment works is capable of producing 34 million litres of water a day.

The pilot plant itself receives a total of 1 m^3/hr, which was split equally between two vessels, one containing a bed of IEX resin, the other the GAC, a schematic is shown in Figure 4.1. The IEX vessel is regenerated every 24 hours using the following scheme:

i) A preliminary slow rinse with permeate water for 5 minutes.

ii) A 10 minute backwash again using permeate water.

iii) A slow rinse for 50 minutes using a NaCl brine solution.

iv) rapid rinsing for 10 minutes using permeate water.

v) A final slow rinse to allow settling of the IEX resin.

The GAC vessel was also backwashed every 24 hours for 10 minutes using permeate water to prevent fouling.

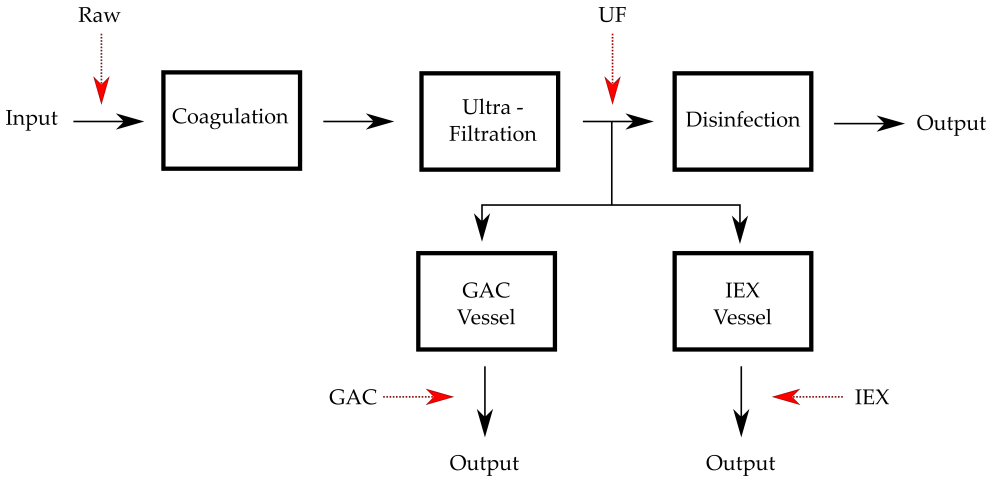


Figure 4.1: Schematic highlighting main processes occurring throughout the pilot plant. Red arrows indicate the four sampling locations.

4.2 EXPERIMENTAL METHODS

Samples were taken once a month for 6 months, beginning March 2019, in order to gain a more representative description on how the system is performing. Samples were collected in acid washed amber glass bottles, sampling locations are indicated on the schematic and volumes are as follows: 2.5 L of raw water, 5 L of UF water and 10 L each for GAC and IEX treated waters. TOC measurements were taken by Scottish Water at the point and time of sampling, these are presented below in Table 4.1. The IEX resin used was the DWI approved resin BW341817 (Barr and Wray), a proprietary anionic resin developed for the purpose of removing DOM from the water, utilised in chloride form. The GAC vessel consisted of 100% activated carbon (Filtec).

(-) ESI FT-ICR-MS analysis was performed using the same method described in Chapter 3. For the ^{1}H NMR analysis, 1mg of sample was dissolved in 500 μl of 99.9% D_2O. Spectra were acquired on an Avance NEO 800 MHz Bruker spectrometer equipped with a 5 mm TCI cryoprobe. Spectra were acquired using the NOESY based water suppression technique as reported in Chapter 3.

4.3 RESULTS AND DISCUSSION

4.3.1 *TOC Data*

Table 4.1: TOC values of the collected samples in mg/L.

Month	Raw	UF	GAC	IEX
February	2.3	1.5	1	0.3
March	2.6	1.8	1.5	0.5
April	2.5	*na*	*na*	*na*
May	3.1	1.8	1.4	1
June	2.3	1.9	1.4	0.2
July	2.5	1.8	1.4	0.4

The TOC values in Table 4.1 highlight that the UF water appears to be relatively stable, with a standard deviation of 7.7%, and that the IEX treatment process is removing most of the organic material, outperforming the GAC treatment significantly every month.

4.3.2 *Overview of FT-ICR-MS Data*

Like in Chapter 3, samples here were analysed using ESI (-) mode ionisation, masses from the same range as previously were used, 200 - 600 m/z. Molecular formulae were assigned using Formularity[121], with molecular constraints of $^{12}C_{0-66}$ $^{1}H_{0-126}$ and $^{16}O_{0-30}$, nitrogen was omitted due to the increased resolution needed. Samples were analysed in duplicate and those formulae which were present in both replicates are presented in Table 4.2.

Table 4.2: Monoisotopic molecular formulae assignments for samples. Formulae presented appeared in both instrumental duplicates.

	Monoisotopic Assignments			
Month	Raw	UF	GAC	IEX
February	2203	2067	1991	1639
March	2061	1924	1957	1553
April	2022	1974	1686	1596
May	2160	2105	2031	1758
June	2157	2078	1765	1458
July	2040	1882	1656	1667

Table 4.2 shows that:

i) The IEX is removing more species overall than any other treatment process, apart from in the month of July, where its removal performance seems to have been matched by the GAC filtration.

ii) The molecular formula count of the raw water appears to be relatively stable, with a standard deviation of only 3%.

iii) The GAC filtrated samples have the highest variation with a standard deviation of over 8%, the highest of all sample sets.

iv) The UF for most months has a similar formula count to that of the raw water, implying minimal changes to the DOM composition.

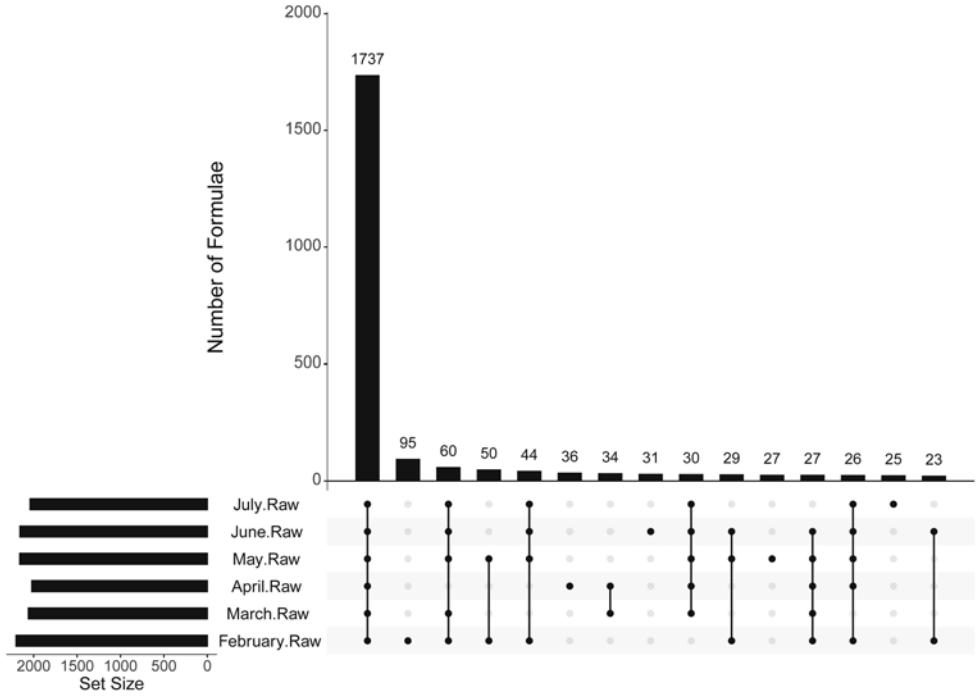


Figure 4.2: UpSet plot of the 6 raw water samples. Each sample represents a pair of instrumental duplicates.

UpSet plots were made for each stage of treatment across the 6 months in order to determine if the composition at each time point was similar or not. Figure 4.2 shows an UpSet plot which highlights the high degree of similarity between the raw water samples. UpSet plots for the remaining sample stages can be seen in Figures B.1, B.2 and B.3. The samples don't show much variation across the 6 month period, with 84% of the total formulae assigned to raw water samples being present in at least 3 of them, the UF, GAC and IEX samples have 86%, 85% and 78% respectively. These formulae present in at least 3 of the 6 samples were used to create representative samples for each treatment stage. Working with representative samples will allow any small monthly variations to be discarded, whilst still retaining the relative instability of the GAC sample, where 3 of the samples appear to have higher assignment numbers. Having representative samples will also make it much easier to compare treatment stages between each other.

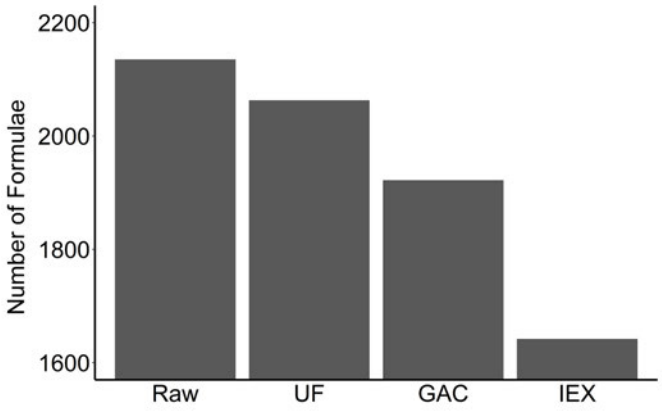


Figure 4.3: Total molecular formulae counts in the representative MS samples.

Figure 4.3 indicates the number of molecular formulae in these new representative samples, showing a slight reduction in molecular formulae after ultrafiltration. This Figure shows that there is merit in the use of a post coagulation technique as both the GAC and IEX have fewer assigned species than the UF treated water. Of the two post coagulation techniques the IEX treatment is not only removing more TOC, but also reducing the organic complexity of the DOM more than the GAC treatment.

4.3.3 ^{1}H *NMR analysis of the treated samples*

These same 24 samples (6 months x 4 treatments) were also analysed using ^{1}H NMR spectroscopy. Like in Chapter 3, the proton spectra were integrated into regions of: aromatic species (6 - 10 ppm), an unsaturated region (4.8 - 6 ppm), carbohydrate region (3.1 - 4.6 ppm), CRAM region (2 - 3.1 ppm) and finally an aliphatic region (0.5 - 2 ppm). Average integral intensities were calculated across the 6-month sampling period for these 5 spectral regions and are presented in Figure 4.4.

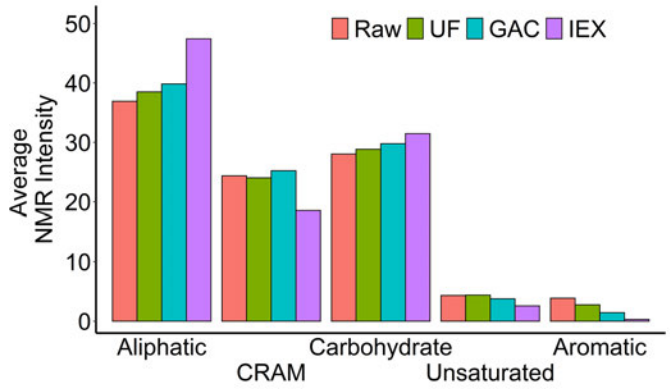


Figure 4.4: Average relative ^{1}H NMR intensities of each integral region for the four samples, Raw, UF, GAC and IEX.

It can be seen that the largest changes occurred for the aromatic molecules. This region is compared in more detail in Figure 4.5, where the relative intensity of the aromatic region across the 6 month sampling period is presented. Equivalent plots for the aliphatic, CRAM, carbohydrate and unsaturated regions can be found in appendix Figures B.4a - d. A clear reduction of aromatic proton intensity is seen for all three treatments, UF, GAC and IEX, in comparison to the raw water.

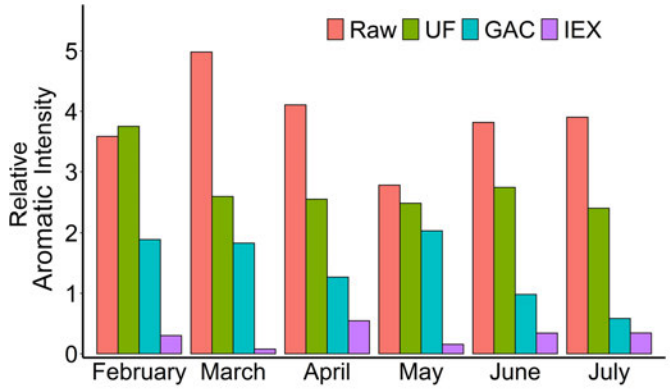


Figure 4.5: Relative ^{1}H NMR intensities of the aromatic region (6 - 10 ppm) for raw, UF, GAC and IEX samples between February and July.

Overall, the following conclusions can be drawn from the NMR data:

i) The relative presence of aliphatic compounds remains relatively constant for the Raw, UF and GAC treated samples (except GAC in June, where intensity is higher), whereas the IEX samples have consistently greater aliphatic content.

ii) The proportion of CRAM molecules is similar in the Raw, UF and GAC samples, but significantly decreased in the IEX samples.

iii) The relative presence of carbohydrates steadily increases as per the following order Raw < UF < GAC < IEX.

iv) The aromatic compounds are gradually depleted by each treatment, in the following order: Raw > UF > GAC > IEX. The concentration of aromatic compounds is significantly lower after IEX treatment every month.

v) The unsaturated compounds follow the trend of aromatic compounds, but their depletion is less severe.

vi) Of the two post coagulation treatment, the IEX results in the more dramatic compositional changes. The GAC treated samples have a similar composition to the UF treated samples.

4.3.4 *MS analysis of the ultrafiltration treated (UF) water*

In order to effectively compare the two post coagulation treatments IEX and GAC, it is necessary to first characterise the water which is fed to the pilot plant, the UF water. This was done by comparing the representative UF sample to that of the raw water.

The FT-ICR-MS results suggest that these two samples are highly similar; they share 2007 molecular formulae with only 128 species being removed by the UF treatment, while 56 new species were introduced. A DBE plot and oxygen series plot based on the representative molecular formulae of the raw and UF data (Figure 4.6), were produced to assess the species removed. These plots suggest that the raw water is slightly richer in lower DBE species and lower oxygen containing species. The UF is thus removing more aliphatic like species or low aromaticity, low oxygen content species. This is a surprising result considering the reduction in aromatic ^{1}H intensity seen in Figure 4.5; it was expected that a majority of the species removed would be of an aromatic nature. This suggests either that the aromatic species being removed by the UF are of a much higher mass than that analysed by the FT-ICR-MS or if they are low mass, (-) ESI does not efficiently ionise them (outcompeted).[122]

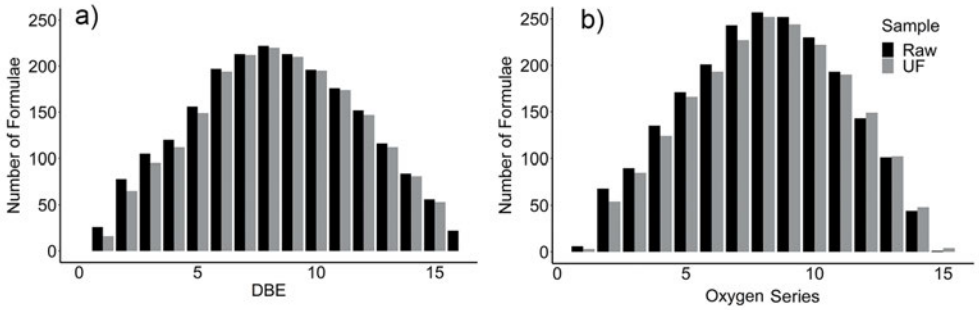


Figure 4.6: Comparison between the raw (black) and UF (grey) treated samples. a) a DBE plot, b) an oxygen series plot.

A closer inspection of the 128 species removed and 56 only present after UF revealed that the most significant difference between these two pools was the molecular weight, shown in Figure 4.7.

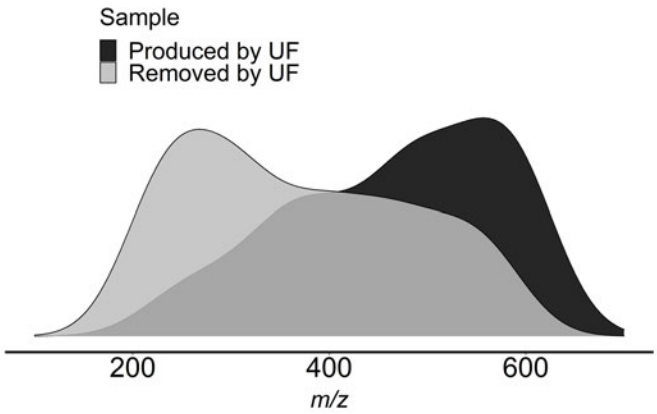


Figure 4.7: Mass density plots for the formulae present only after UF treatment (black) and those removed by UF (grey). Samples are scaled individually.

This is a surprising result as coagulation followed by filtration has been shown to preferentially remove larger molecular species,[122] whereas Figure 4.7 highlights that the majority of those removed by UF (grey) are of low molecular weight. Those species which are introduced following UF appear to be a larger molecular weight, although this is likely an artefact of competitive ionisation.

AI_{mod} plots were also produced to examine the composition of the UF treated sample to that of the raw water (Figure 4.8). Compounds were designated as either polycyclic aromatics (group I, $AI > 0.66$), plant-derived polyphenols (group II, $0.5 < AI \leq 0.66$), highly unsaturated and phenolic compounds (group III, $AI \leq 0.5$ and $H/C < 1.5$) and aliphatic compounds (group IV, $AI \leq 0.5$ and $H/C \geq 1.5$) as per Kellermans modified classification.[104] There is very little difference in group composition with the largest percentage difference being

observed for group I. Here the UF formulae are reduced to 75% that of the raw water, groups II and IV are approximately 92% while group III is 101%. These results indicate that no compound class is being strongly targeted but that there is a slight preference in the removal of group I species, and no removal of group III species, the slight gain of species in this group could be explained by the effects of competitive ionisation.

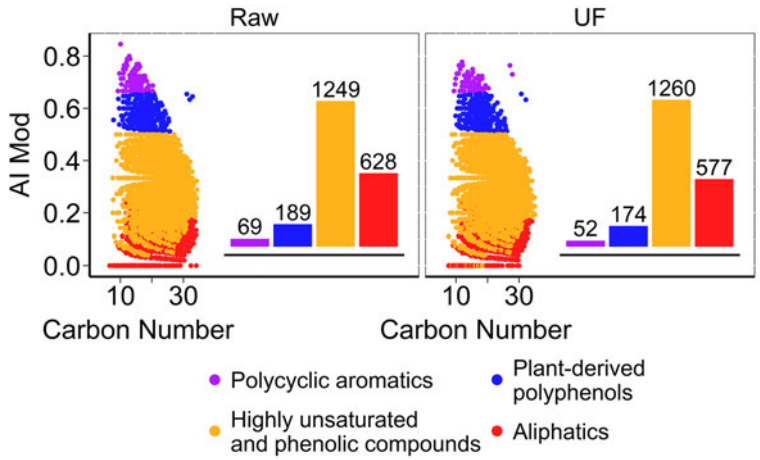


Figure 4.8: AI_{mod} plot comparing molecular formulae assigned in the representative raw and UF samples.

The FT-ICR-MS shows that the coagulation and ultrafiltration are causing minimal changes to the complexity of the DOM, the ^{1}H NMR data shows the UF sample has a slight increase in aliphatic compounds and a slight decrease in aromatic species, but generally corroborates the ICR-MS data. Despite the fact that the composition appears relatively unchanged, the TOC for the UF treated water is consistently lower than the raw water, suggesting that the DOM is being removed by this treatment in a non-selective manner.

4.3.5 *Comparison between representative GAC and IEX samples*

The goal of this study was to determine whether GAC or IEX filtration performed better at removing organics when implemented as a post coagulation treatment. A better post coagulation treatment will result in not only a greater reduction of overall organics (TOC), but also a greater reduction in those species which are likely to go on and react with a disinfectant to produce DBPs

In this section, the representative samples of the GAC and IEX treatments will be directly compared to the UF water in order to determine if the treatment processes are resulting in the removal of compounds with particular characteristics. The section following this, will investigate the two post coagulation

treatments during the month of July when the GAC treatment was performing at its best (highest level of molecular formulae removal).

It has already been mentioned, but it should be reiterated, the TOC values reported in Table 4.1 show that the IEX consistently outperforms the GAC treatment at reducing the overall concentration of DOM in the coagulated water. Also the number of monoisotopic formulae assignments (Table 4.2, obtained from the MS data, is lower every month for the IEX treatment other than July, where the values are comparable (1656:1667, GAC:IEX).

The representative samples obtained for the GAC and IEX were compared to the representative sample produced from the UF data, to provide insight into the molecules that these two post coagulation treatments are targeting. The formulae present in the UF sample that are no longer present in either the GAC or IEX were designated as removed, those that were not present in the UF sample but are in the GAC or IEX samples were designated as being produced by the treatment process. It should be emphasised that there are several ways in which these newly appeared formulae could be introduced, whether it be a genuine introduction from the treatment or that these newly present species are just no longer being outcompeted at ionisation due to the removal of other highly ionisable species. The numbers of produced and removed formulae are presented in Table 4.3.

Table 4.3: Formula numbers of compounds produced/removed by the GAC and IEX treatments

Formula	GAC	IEX
Removed	229	593
Produced	88	172

Unsurprisingly, Table 4.3 shows that the IEX treatment has removed more species, 593 vs 229 of the GAC. These removed formulae will be investigated to describe the type of species each of these techniques are targeting. A DBE plot comparing these two sets of formulae is presented in Figure 4.9a. This figure shows that both of the treatment stages appear to be preferentially targeting compounds with a relatively high DBE, the maxima being between 12 and 14 for the IEX and GAC respectively, compared with DBE maximum of 8 for the UF samples (see Figure 4.6a). Interestingly, those removed by GAC appear to have a second maximum at 5 DBE. However, Figure 4.9b, arguably more interesting, shows an oxygen series plot which indicates that the IEX is preferentially removing species with relatively high numbers of oxygen, whereas the species removed by GAC treatment have a maximum coinciding with that of

the UF water (O_8). As the IEX treatment relies upon ionic interaction between the resin and the organic species, it is logical that species with multiple acidic sites, i.e. more oxygens are more likely to bind and eventually be removed.

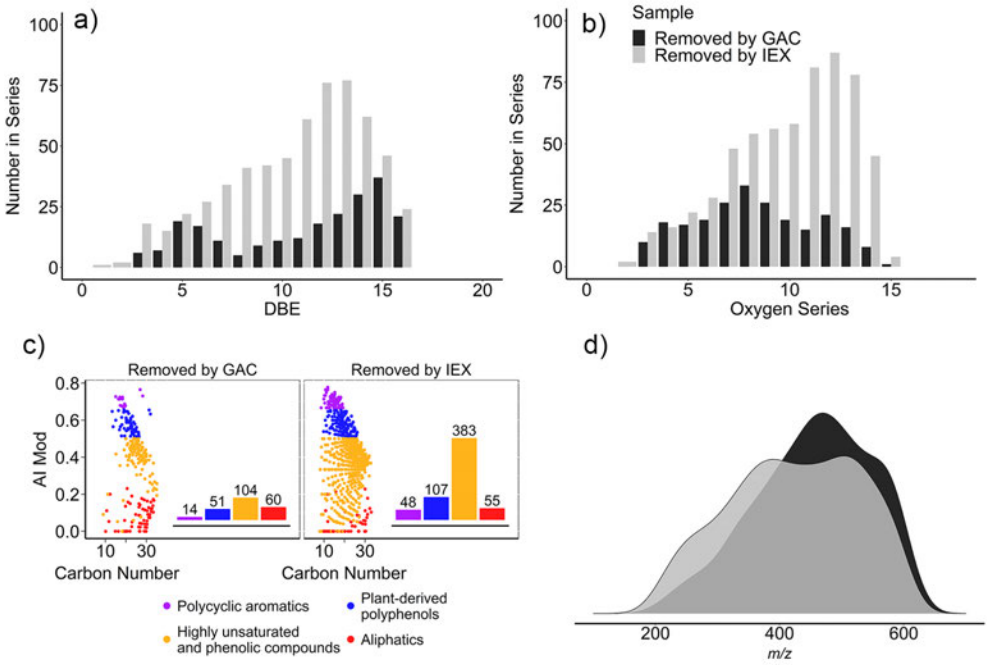


Figure 4.9: A a) DBE plot, b) oxygen series plot, c) AI_{mod} plot and d) mass density plot of the formula removed from the UF water by the GAC treatment (black) and the IEX treatment (grey).

The mass density plot (Figure 4.9d) highlights that both treatments are removing species across the entire mass range analysed with the IEX treatment having a stronger presence at lower M_W values, whereas the GAC treatment appears to have a tendency to remove larger molecular species.

AI_{mod} plots produced (Figure 4.9c) to compare compound class distributions, highlight that with both treatments, the majority of the compounds removed are characterised as highly unsaturated phenolics, however that is likely because this class was the most populated to begin with.

Table 4.4: Assignment numbers of compounds removed by GAC and IEX treatments, based on the Kellerman AI_{mod} classification. Numbers in brackets indicates percentage removal based on the UF treated representative sample.

Group	GAC	IEX	IEX/GAC
I	14 (26.9)	48 (92.3)	3.4
II	51 (29.3)	107 (61.5)	2.1
III	104 (8.3)	383 (30.4)	3.7
IV	60 (10.4)	55 (9.5)	0.9

Therefore, a more informative comparison of the relative removal efficiency of the two treatments is presented in Table 4.4. It shows that the IEX treatment resulted in over a 92% removal of group I species from the UF treated water, whereas the GAC treatment only removed 27%, i.e. the GAC is 3.4 times less efficient. This implies that the IEX treatment has a high selectivity for group I species, i.e. those with a high AI_{mod} and DBE (the latter can be seen in Figure 4.9a). Class II and III compounds are removed by IEX 2.1 and 3.7 times more efficiently, while the GAC treatment removed the aliphatic compounds slightly better.

The NMR data presented in Figure 4.5 also shows that the IEX treatment outperforms the GAC, with the former resulting in less relative signal intensity in the aromatic region, a region highlighted as it is thought that more conjugated species result in the production of more DBPs.[123] Conversely, this drop in aromatic species results in a relative increase of aliphatic species, shown in Figure 4.4.

To summarise the removed formulae as implied from the analysis of representative samples, both treatments are removing species characterised by a relatively high DBE, both preferentially remove group I and II species, i.e. those which are highly unsaturated. The biggest difference however is in the oxygen content of the removed species, with the GAC treatment preferring those with the average numbers of oxygen, the IEX preferring those with the average numbers of oxygen, the IEX preferring high numbers of oxygen. The removal of aliphatic compounds is equally inefficient, while the IEX removed on average 2.5 times more compounds than GAC.

The gained compounds are more difficult to interpret, as the method by which they are introduced is not clear. Overall, the gained molecular formulae constitute 4.3 and 8.3% of the formulae identified in the representative UF sample set. In both GAC and IEX treatments the aliphatics represent the largest number of new formulae. In the GAC sample these represent the largest

proportion relative to the original amount in this class (around 10%), while in the IEX sample classes I and II registered an increase of approximately 30%. Their DBE, oxygen series and mass density plots are presented in Figure 4.10.

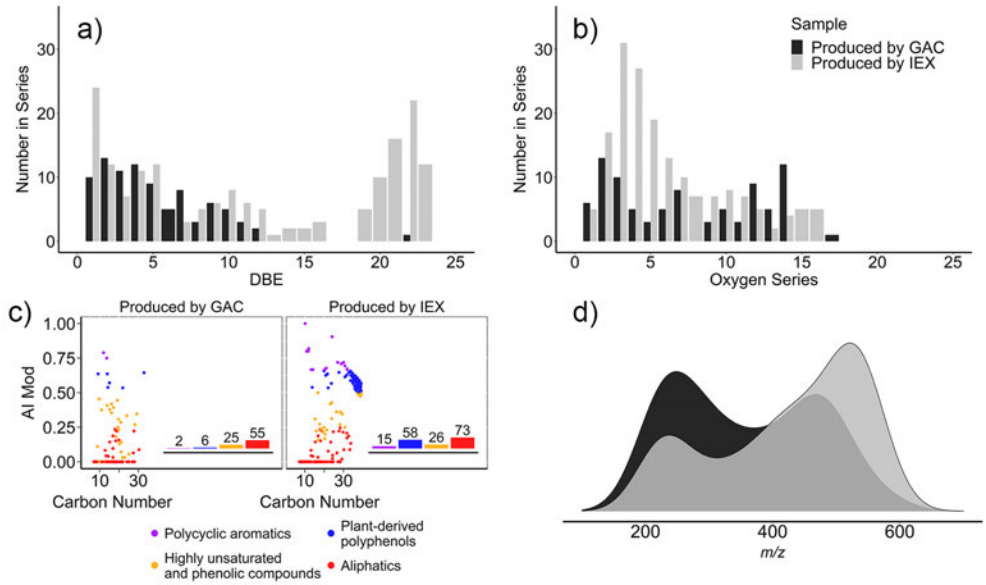


Figure 4.10: a) A DBE plot, b) Oxygen series plot, c) AI_{mod} plot and d) a mass density plot of the newly gained molecular formulae produced by the GAC treatment (black) and IEX treatment (grey) vs the UF representative sample. The inset bar graph on the AI_{mod} plot indicates the count in individual compound classes.

4.3.6 *Comparison between GAC and IEX July samples*

Table 4.2 shows that the GAC assignment numbers are not as consistent as the other samples with a standard deviation across the 6 months twice as large as for the other three sample types. This section will therefore compare the characteristics of the GAC and IEX treatment during July, the month where the GAC treatment resulted in the greatest reduction of assigned formulae.

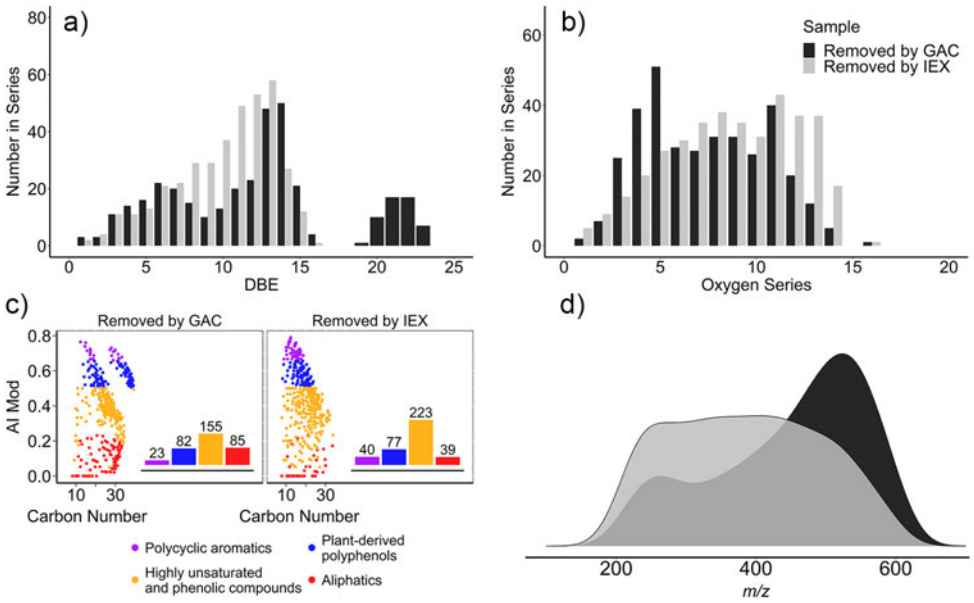


Figure 4.11: a) A DBE plot, b) Oxygen series plot, c) AI_{mod} plot and d) a mass density plot of the molecular formulae removed by the GAC treatment (black) and IEX treatment (grey) vs the UF July sample. The inset bar graph on the AI_{mod} plot indicates the count in individual compound classes.

By analysing the characteristics of the species removed by GAC and IEX throughout July we can conclude that:

i) The GAC treatment during this month removed a cluster of high DBE compounds that the IEX treatment did not. These compounds have a DBE > 20.

ii) The formulae removed by the GAC treatment are larger on average than those removed by the IEX.

iii) The oxygen content of the removed compounds are similar between the two treatments, but like the representative 6-month samples, the GAC treatment has slight preference for lower oxygen content species.

iv) From the AI_{mod} plot in Figure 4.11c, we can see that again the GAC treatment appears to be more selective than the IEX towards the removal of the aliphatic compounds.

v) The IEX treatment again is still better at removing group I species, those which have the highest AI_{mod} values.

In this month, where both GAC and IEX treatments resulted in the removal of similar formulae numbers, the characteristics of each treatments removal

remains similar to those observed in the representative samples. The only difference in July is the cluster of high DBE compounds removed by the GAC treatment, which is pushing the mass distribution to higher M_W values.

4.4 CONCLUSIONS

This chapter has compared two post coagulation treatments in their DOM removal characteristics and found that the IEX treatment is a much more stable treatment system, with the GAC resulting in the greatest formulae assignment variation of the four samples studied. The IEX treatment consistently resulted in a greater reduction of aromatic compounds, being lower in all 6 months. The IEX treatment results in the removal of high oxygen containing species, which is consistent with the removal mechanism, i.e. the more acidic groups a molecule has, the more likely it is to exchange with the IEX resin. Taking these characteristics into account, the IEX treatment is better in terms of overall TOC removal, but also the left over DOM has more favourable characteristics and is less likely to react with disinfectants like chlorine producing less DBPs.

Part III

PHOTOCATALYTIC DEGRADATION OF DISSOLVED ORGANIC MATTER

PHOTOCATALYTIC TREATMENT OF SRFA: CATALYST COMPARISON

5.1 DECLARATION

This chapter was a joint effort between Victoria Porley and myself.

Victoria contributed to experimental design, the running of the experiment itself, the preparation of the catalysts and to the acquisition/analysis of UV data. I was involved in the experimental design, running of the experiment and acquisition/interpretation of UV, NMR and MS data.

5.2 OVERVIEW

As discussed within the introduction of this thesis, there exists a demand for rigorous treatment processes capable of removing not only those organics that are produced via anthropogenic activities, but also capable of significantly reducing the level of DOM within the potable water supply line. Advanced Oxidation Processes (AOPs) have been touted as being the water treatment processes of the 21st century as they are capable of significantly reducing TOC.

Heterogeneous photocatalysis has received a lot of attention with regards to water treatment applications, however, the majority of groups when investigating a new process or catalyst still use low resolution techniques, such as UV absorption spectroscopy, focusing on the degradation of a single pollutant or a very small mixture (3-5 compounds). Given that the mechanism of degradation likely requires the target molecule to adsorb onto the surface of the catalyst, it is important not only to analyse a single molecule, but the entire organic profile (as much as possible) of the water. However, should complete mineralisation of the molecule not occur, it is also important to understand the molecular composition of any by-products that result.

This chapter uses the complex mixture Suwannee River Fulvic Acid (SRFA), a DOM standard obtained from the International Humic Substances Society (IHSS), to investigate the photocatalytic degradation of DOM using multiple photocatalysts. The catalysts used throughout this study are pristine TiO_2 and a TiO_2 based catalyst which has been modified with bismuth titanate ($Bi_4Ti_3O_{12}$), herein referred to as BTO - TiO_2. The BTO - TiO_2 catalyst

has been shown to possess significant improvements over pristine TiO_2 in the degradation of a variety of different pollutants, such as pesticides, drugs and explosives.[124,125] This improvement is due to the smaller band gap of bismuth titanate (2.96 eV, 419 nm) in comparison to that of pristine TiO_2 (3.2 eV, 387 nm), meaning the bismuth titanate can utilise a larger range of wavelengths. As the conduction and valence band edges are at different locations in the hybrid material a heterojunction is created which acts to reduce the relaxation of excited charges preventing the recombination of holes and electrons, ultimately making the hybrid photocatalyst more efficient.

5.3 EXPERIMENTAL METHODS

Throughout this experiment, the two catalysts being investigated, pristine TiO_2 and BTO - TiO_2, have been immobilised onto soda-lime glass beads, their production has been previously reported.[124] A monolayer of beads, with an average weight of 21 g, was used for each iteration of the experiment. Each bead weighs 0.04 g and has a catalyst loading of 0.2 mg per bead. For each experimental run, 25 mL of stock SRFA solution was used, providing a total mass of 4.2 g/L of immobilised catalyst. This particular SRFA standard (Standard II - *2S101F*) is reported as having 52.34% carbon,[126] ultimately, 22 mg/L of SRFA was chosen to provide a TOC of approximately 11 mgC/L, which is not uncommon, at least in the UK. The experiment was conducted within a 70 x 40 mm glass crucible, wherein the glass beads were completely submerged by the sample solution, which was irradiated by an LED from 5 cm above the surface of the solution. The entire apparatus was enclosed in a light proof surround housed within a fume hood, at room temperature without stirring. A schematic is shown in Figure 5.1.

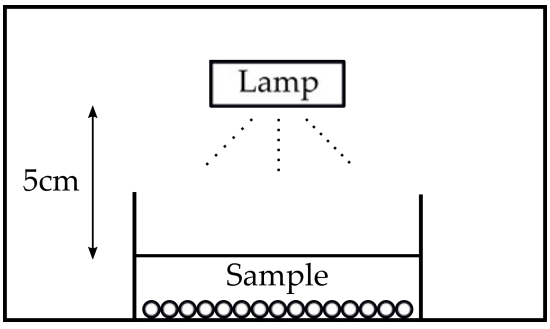


Figure 5.1: Schematic of experiment setup. Sample is contained within a glass crucible with a monolayer of coated glass beads, UV lamp is consistently housed 5 cm above the sample surface, entire setup is self contained.

Three LED's were used throughout the experiment, a 370 nm LED, a 410 nm LED and a white LED, all of which were run at 4W and sourced from Intelligent LED Solutions. Emission spectra of the LEDs are shown in figure 5.2, where it can be seen that the white light LED had a distinct emission band at 430 nm. The entire experiment was run for 3 hours, with samples being taken for FT-ICR-MS and NMR at each hour point, samples were taken at every half hour point for UV analysis. The samples taken for FT-ICR-MS and NMR were lyophilised in order to concentrate the organic matter. There was no need to perform RO and electrodialysis, due to the small sample volumes and the fact that SRFA has already been desalted.

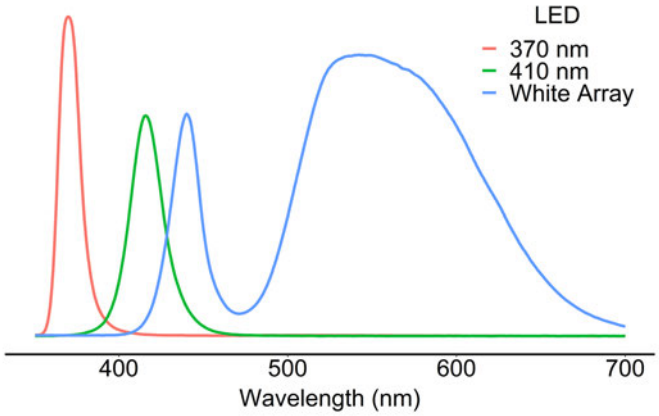


Figure 5.2: Emission spectra of the 370 nm, 410 nm and white LED's used in the photocatalysis experiment.

Like in Chapters 3 and 4, samples were analysed using ESI (-) mode ionisation, masses between 150 and 600 m/z were analysed if they had a SNR ≥ 7. Molecular formulae were assigned using Formularity,[121] with molecular constraints of $^{12}C_{0-66}$ $^{1}H_{0-126}$ and $^{16}O_{0-30}$, nitrogen was again omitted due to the increased resolution needed as explained before. Spectra were acquired using the same instrumental setup reported in 3.

For NMR analysis of the photocatalysis samples, lyophilised DOM was dissolved in 99.9% D_2O as before, however the concentration remains unknown, as the material was too light to accurately weigh. Spectra were acquired on an Avance Neo 800 MHz Bruker spectrometer equipped with a 5mm TCI cryoprobe. Spectra were acquired using the NOESY based water suppression technique with the relaxation times and acquisition times being set to 9 and 2 seconds respectively. 128 scans were acquired per spectrum before being zero filled to 128k points. An exponential line broadening of 1 Hz was applied prior to Fourier transformation. ^{1}H NMR spectra were integrated using 0.01 ppm bins ranging from 0 - 9 ppm and each spectrum was scaled separately as I_i / $\Sigma(I_i)$ where

I_i is the intensity at a given integral. PCA was performed on these integrals using the R package pcaMethods (version - 1.80.0) with pareto scaling applied.

5.4 RESULTS AND DISCUSSION

5.4.1 *Photolysis Experiments*

Throughout this study, experiments were performed where the SRFA solution was irradiated with each LED for the 1, 2 and 3 hours, in the absence of a catalyst, to determine if there was a significant impact from photolysis. As can be seen from the UV data presented in Figure C.1, there does not appear to be any significant changes occurring over time regardless of the LED used, the slight fluctuations seen here were deemed to be from evaporation over the course of the experiment.

An UpSet plot was also produced of the MS data obtained from this experiment, presented in Figure 5.3, this shows that there are 1926 molecular formulae common to all of these blank samples and the SRFA sample. This high similarity shows that irradiation with the LED's alone, did not produce any appreciable difference in the composition of the SRFA.

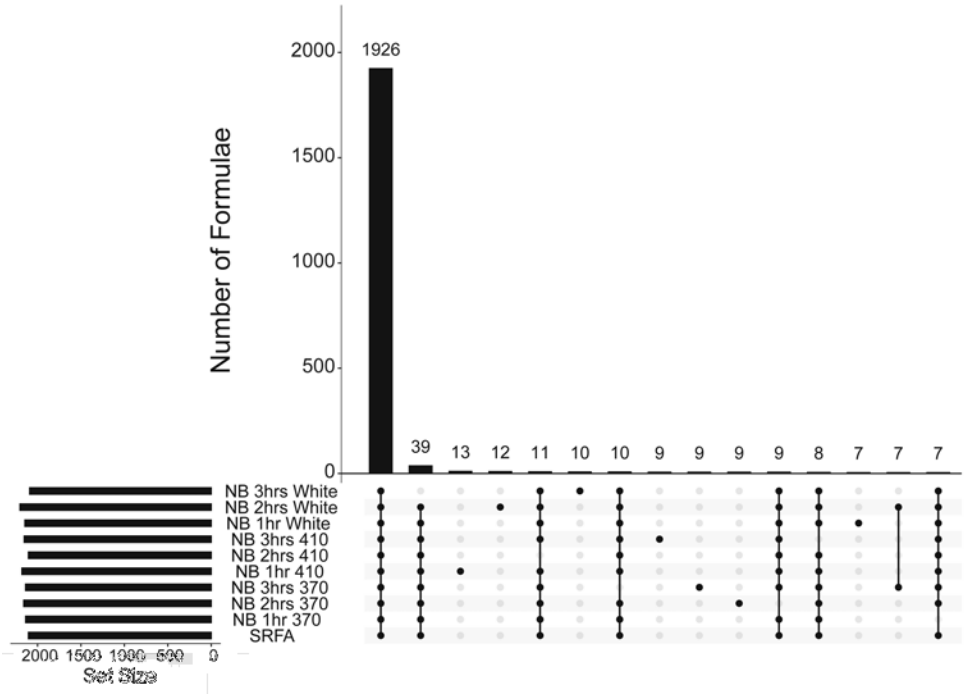


Figure 5.3: UpSet plot of molecular formulae assigned in all blank samples (No Beads, NB) and SRFA. Each blank sample is represented by two experimental duplicates.

After deciding that irradiation in the absence of the catalyst had no observable effect, molecular formulae which were identified in at least 9 of the 10 samples presented in Figure 5.3 were combined to create a composite blank sample. This blank sample contains 2020 molecular formulae and is used throughout this chapter to compare with other sample sets.

5.4.2 *Initial overview of the MS and UV data*

5.4.2.1 *MS Data*

The numbers of molecular formulae assigned, lost and produced by the irradiation across 18 samples, that is three time points across three LEDs and two catalysts, were calculated by comparing these samples to the composite blank, they are presented in Figure 5.4.

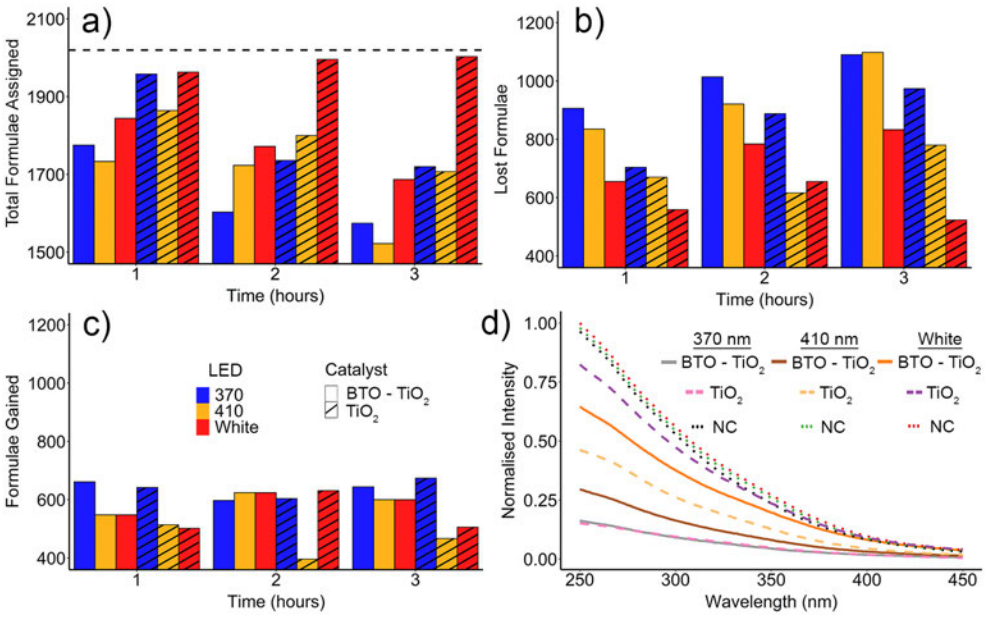


Figure 5.4: The count of molecular formulae assigned in samples containing the BTO - TiO_2 or TiO_2 catalysts, irradiated with the 370 nm, 410 nm or the white LED for 1, 2 and 3 hours. a) Total molecular formulae assigned, b) formulae lost, c) formulae produced by irradiation and d) normalised UV absorption intensity after 3 hours of irradiation relative to the blank experiments. The dashed line in a) represents the 2020 compounds of the composite blank sample.

Figure 5.4 highlights that with the exception of the white LED irradiation in the presence of the TiO_2 catalyst, the total number of molecular formulae identified gradually decreases from 1 to 3 hours. The decrease is not dramatic,

with the BTO - TiO_2 sample, irradiated with the 410 nm LED losing only 25% of formulae after 3 hours, compared to the composite blank sample (see Figure 5.4a). However, this measure is ultimately flawed and masks the true dynamics of the system, which are revealed by analysing Figures 5.4b and 5.4c, which show the molecular formulae lost and those which are newly identified respectively. Up to 54% of the initial molecular formulae were degraded resulting in additional formulae, degradation by-products, up to 32% of the initial value. It can also be seen that the majority of the changes in all experiments occurred within the first hour, with the total molecular formulae reducing by approximately 10% when irradiated for a further 2 hours. There also appear to be appreciable differences between the two catalysts, with the BTO - TiO_2 catalyst resulting in the loss of more original compounds at every time point and wavelength, when compared to the pristine TiO_2. In contrast, the number of produced by-products appears to be comparable for not only the catalysts, but also the wavelengths used to irradiate and the length of irradiation.

5.4.2.2 *UV Data*

Monitoring the photocatalytic degradation using normalised UV absorbance at 250 nm (A_{250}), after three hours of irradiation (Figure 5.4d) shows that:

i) The photolysis was marginal in the absence of a catalyst, however, a very slight increase in absorbtion is observed as the wavelength decreases.

ii) When the catalysts are present, the reduction in UV active compounds increases substantially as the wavelength of irradiation is decreased, an observation that correlates well with the MS data.

iii) The photocatalytic process was comparable for both catalysts at 370 nm, resulting in an absorbance reduction of 82% that o the starting value.

iv) When irradiated with either the white or 410 nm LED, the BTO - TiO_2 performed more efficiently than the pristine TiO_2.

As is highlighted in a study by Huang et al,[127] the UV data obtained from a complex mixture cannot be used as a proxy for the measurement of DOM concentration for a multitude of reasons. The number of compounds absorbing at the monitored wavelengths is not known, there are unknown differences in their extinction coefficients and finally the system is highly dynamic, with the concentration of compounds changing unequally over time. Despite this, interpretation of the time series UV data using pseudo-first order kinetics is

still valid and can be used to rank the rates of photocatalytic degradation of UV_{250} active compounds.[127]

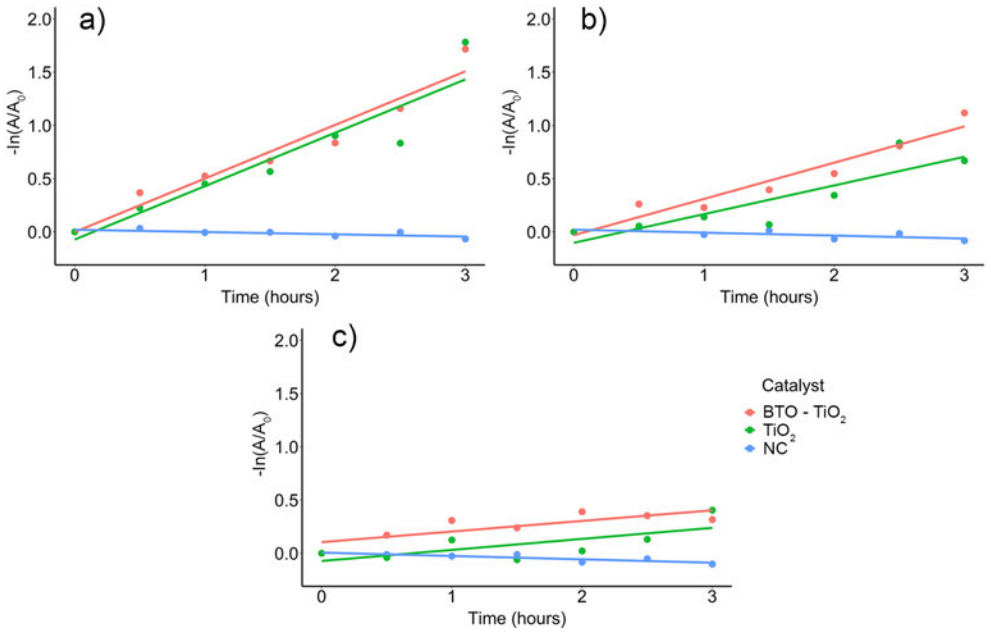


Figure 5.5: Pseudo-first-order rates at a) 370 nm, b) 410 nm and c) white LEDs and individual catalysts (NC - no catalyst) based on the UV_{250}

As can be seen by the rate plots in Figure 5.5, the rate of degradation is linked to the wavelength of irradiation as follows 370 nm > 410 nm > white LED, again the sample with no catalyst appears to slightly increase in absorption, this may be caused by evaporation. At it's greatest there is an 82% reduction in UV_{250} absorbance, when compared to a 54% loss of molecular formulae assignments (not counting an additional 32% new formulae) suggesting that either (i) the UV active compounds are being targeted disproportionally, (ii) that the strongest UV absorbing compounds are too large to be observed via MS, (iii) or their characteristics do not lend themselves to ionisation via (-) ESI.

An existing study monitored the photocatalytic degradation of peat DOM by measuring the UV absorbance, the TOC content and by FT-ICR-MS, like here, they also show a similar dramatic decrease in UV absorbance, while the TOC levels and number of assigned molecular formulae decrease at a slower pace.[128] This implies that there are few compounds which are mineralised completely to CO_2 and that there is substantial bleaching of the aromatic species throughout the duration of our experiment.

What follows is a more in-depth analysis of the molecular changes occurring throughout the photocatalysis process, the samples as a whole, the degraded compounds and the by-products formed.

5.4.3 *Molecular characteristics of photocatalytic degradation*

An UpSet plot of the molecular formulae assigned within the 370 nm irradiated BTO - TiO_2 samples and the composite blank sample is shown in Figure 5.6. As was mentioned above and indicated by the set sizes at the bottom left of the figure, the total number of assigned formulae, changes most within the first hour, then very little for the remainder of the experiment.

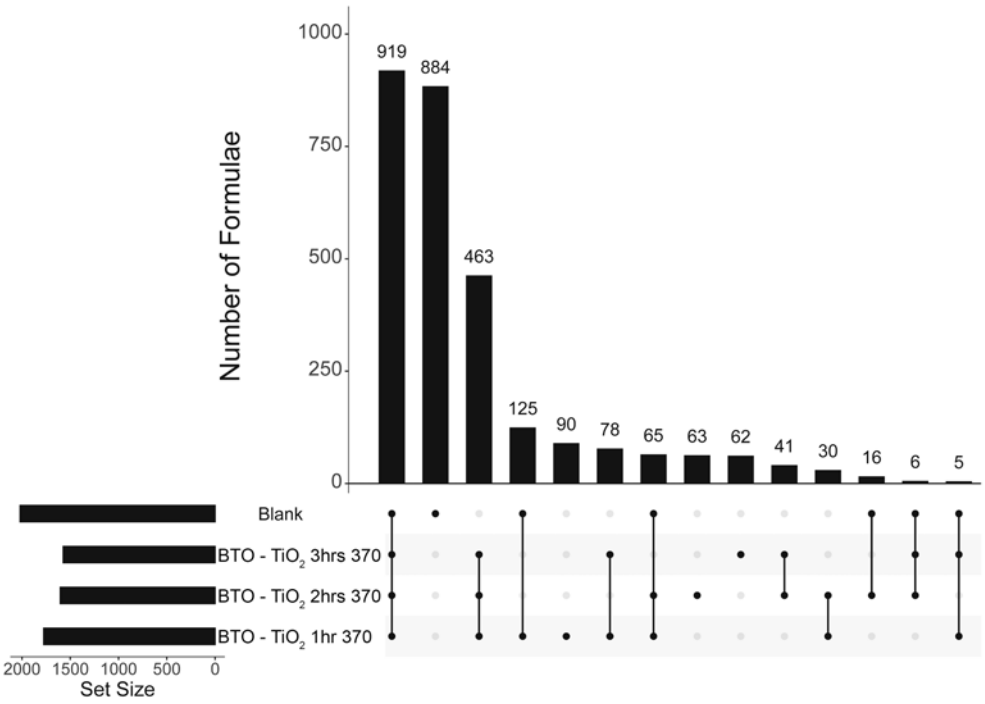


Figure 5.6: UpSet plot of molecular formulae assigned in samples irradiated with the 370 nm LED for 1, 2 and 3 hours in the presence of the BTO - TiO_2 catalyst, including the reference blank data set.

The UpSet plot shows that there are 919 formulae which are not removed completely by the three hours of photocatalysis at 370 nm, as is indicated by their presence at all time points, with or without the presence of a catalyst. The next intersection highlights 884 molecular formulae which are no longer present in the in the irradiated samples (removed or degraded completely); the corresponding compounds were removed quickly, as they were absent already after the first hour of irradiation. The third intersection, containing 463

distinct molecular formulae, represents compounds that are produced by the photocatalysis within the first hour and survive the remainder of the experiment, these are referred to here on as degradation by-products. Beyond these three intersections, there are smaller intersections which represent formulae which are unique to one or multiple time points, they indicate molecular changes beyond the first hour of the experiment. However, as smaller intersections could be the product of fluctuations in peak intensity around the signal to noise limit, not to mention it is impractical to analyse every intersection, only those which represent the majority of molecular changes throughout the course of the experiment are analysed here. Therefore, only the initial 3 intersects, mentioned above, are used for the examination, a choice which is further justified by the fact that the majority of complete compound removal occurred within the first hour of irradiation.

An initial observation is that there are many more molecular formulae lost (884) than there are produced (463). This can be explained multiple ways.

i) A considerable number of species could have been mineralised completely to form CO_2 and H_2O, although this is unlikely given the small extent of mineralisation reported by previous studies on TiO_2 catalysed degradation.[128]

ii) The species could have been degraded into species which were too small for detection with the ICR-MS (> 80 m/z).

iii) Species which were initially different may degrade along common pathways producing by-products which have identical formulae. (e.g. decarboxylation of molecules with a common molecular skeleton differentially decorated by carboxylic groups).

iv) There also exists a possibility that due to electrospray ionisation being a competitive process, some less ionisable molecules, although present from the beginning, became observable only after irradiation, as they were initially outcompeted by those species now lost.

Equivalent UpSet plots were produced for the 410 nm and the white LED (Figures C.2 and C.3). These plots show that the irradiation at 410 nm caused similar changes to the 370 nm samples, whereas irradiation with the white LED resulted in the degradation of approximately 25% less compounds. The number of degradation by-products was comparable across all three LEDs.

Analogous UpSet plots for the TiO_2 catalyst can be found in Figures C.4, C.5 and C.6. Although general trends seen here are similar to those observed for the BTO - TiO_2 catalysts, there are differences that are best visualised

by comparing the preserved, removed and newly detected molecular formulae contained in the first three intersections of all six UpSet plots. Assorted by individual wavelengths, their count is presented in Figure 5.7 for both the BTO - TiO_2 and pristine TiO_2 catalysts.

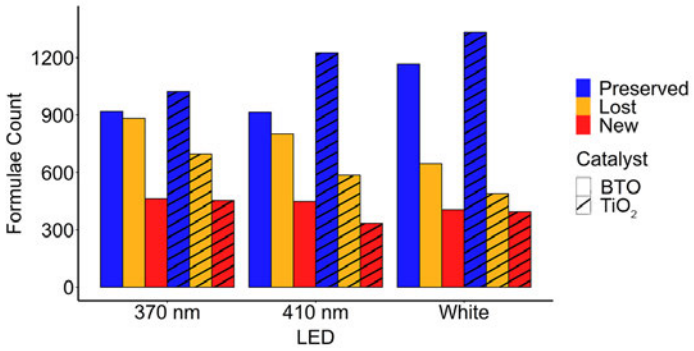


Figure 5.7: The number of preserved, lost or newly detected molecular formulae at individual wavelengths, common to all three time points corresponding to the first three intersections in the respective UpSet plots. The solid bars represent the BTO - TiO_2 samples, the dashed bars represent the TiO_2 samples.

It can be seen that the gap between the preserved and lost formulae is larger for the TiO_2 catalyst at each wavelength, implying a less efficient degradation of species in comparison to the BTO - TiO_2 catalyst. At the same time, the number of lost formulae decreases as the wavelength of irradiation increases for both catalysts. Also, as observed previously when considering the individual time points (Figure 5.4), the number of new formulae (by-products), is approximately constant across all wavelengths and both catalysts.

The degraded compounds and the degradation by-products presented in Figure 5.7 for the BTO - TiO_2 catalyst, were further inspected with the aim to discern the molecular characteristics of the compounds they represent. Towards this end a composite characterisation plot was produced (Figure 5.8), consisting of a a) DBE plot, b) Oxygen series plot, c) an AI_{mod} plot utilising a previously described classification system and d) an m/z density plot.

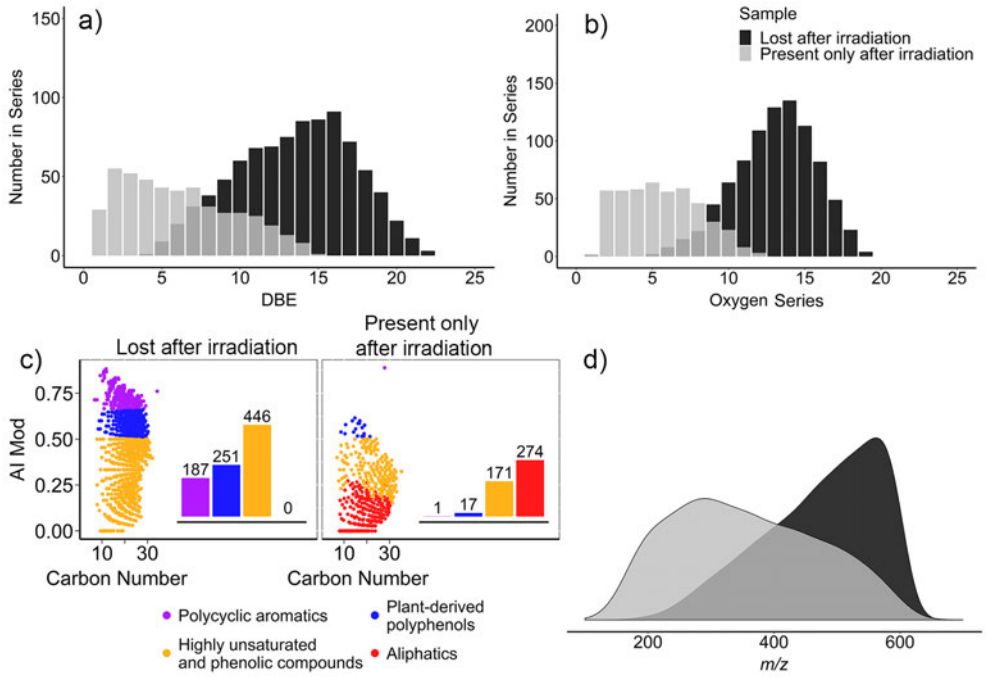


Figure 5.8: Molecular characteristics of the changes to SRFA upon irradiation with a 370 nm LED in the presence of the BTO - TiO_2 catalyst. The degraded compounds (884) are compared to the newly formed degradation by-products (463). a) a DBE plot, b) oxygen series plot, c) AI_{mod} plot, with Kellerman classification and d) a m/z density plot, normalised independently to the sum of all points in each group.

As can be seen by Figure 5.8, the species which are lost have both a higher DBE and oxygen number than those newly produced compounds and there is a noticeable depletion in aromatic species with an $AI_{mod} > 0.5$. The Kellerman classification was employed on the AI_{mod} index, highlighting that none of those compounds which have been degraded are classed as aliphatic. Of the degradation by-products, practically none of them belong to the most aromatic classes, the polyphenols and polycyclic aromatics, while a significant proportion of these products belong to the aliphatic group.

It can also be observed that a high proportion of the molecular species which were degraded during the photocatalysis, have a higher molecular weight than the by-products. Larger molecular weight species are thus not mineralised completely and instead are broken down to form lower molecular weight by-products. A reduction in the molecular weight of DOM species during photocatalytic degradation has been observed previously[129–131], through the use of High-Performance Size Exclusion Chromatography (HPSEC) and UV detectors. The detection method used in these studies is only capable of observing UV active compounds, also it is known that DOM in solution can

form aggregates, consequently the SEC response would reflect the combination of genuine molecular weight changes and the formation/disintegration of these aggregates therefore describing apparent molecular weight changes. In contrast, FT-ICR-MS is capable of monitoring molecular changes to individual molecules regardless of whether they are UV active or not.

The composite characterisation plot shown above for the BTO - TiO_2 sample (Figure 5.8), was replicated for the 410 nm and white LEDs, these are shown in Figures C.7 and C.8 respectively. Analogous plots for the TiO_2 catalyst are also presented in Figures C.9, C.10 and C.11. The conclusions drawn from these plots are similar to those obtained from the 370 nm LED and are reiterated below.

i) The degraded species, on average, have a higher DBE and oxygen count than the degradation by-products.

ii) The degraded species tend to rank higher on the AI_{mod} scale and as a consequence, fall into more aromatic categories as defined by the Kellerman classification.

iii) The produced formulae are on average of a lower mass than those compounds which were degraded, suggesting they are degradation by-products and that there are low levels of complete mineralisation.

In order to obtain a greater understanding between the individual LEDs and catalysts, the compounds which were degraded are analysed in more detail below.

5.4.4 *The degraded compounds*

The formulae removed during photocatalysis were analysed previously in their entirety with reference to the 2020 formulae of the starting material. However, their classification into compound groups, makes a more detailed analysis possible.

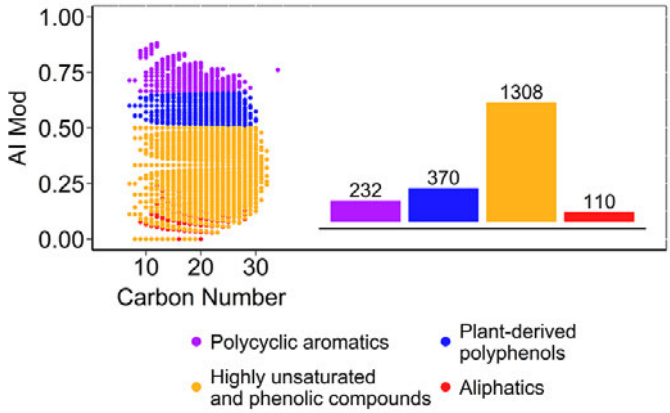


Figure 5.9: An AI_{mod} plot highlighting the molecular characteristics of the 2020 molecular formulae present in the starting material.

This classification revealed an uneven distribution of compounds in the initial DOM sample (Figure 5.9) with 11.5% of them being polycyclic aromatics (**I**), 18.3% plant-derived polyphenols (**II**), 64.8% highly unsaturated and phenolic compounds (**III**) and 5.4% being aliphatic compounds (**IV**). Focusing on the formulae removed in experiments with both catalysts and all three LED sources (Figures 5.8c and C.7 - C.11), changes are quantified next in reference to the number of compounds in groups **I** - **IV** of the starting material.

As expected, the compound degradation is both wavelength and group dependent. For the BTO - TiO_2 catalyst, from the 370 nm to the white LED light source a drop from 80.6% to 56.5% (group **I**), 67.8% to 48.1% (group **II**), and 34.1% to 25.8% (group **III**) in the efficiency of compound removal was observed (Figures S16). This data shows that the different compound classes have different levels of susceptibility, with the polycyclic aromatics (**I**) being the most susceptible followed by the plant-derived polyphenols (**II**) and the phenolic compounds (**III**). Aliphatic compounds are not included in this analysis because the numbers lost are negligible.

Overall there is an approximate reduction of 30% in the efficiency of removal with the white LED vs the 370 nm LED. The same analysis was performed for the pristine TiO_2 catalyst where the trends remain the same, although the efficiency at each wavelength is on average 76% that of the BTO - TiO_2 catalyst.

When analysing the composition of the pools of degraded compounds, it is interesting to note that the first three compound classes (**I** - **III**) are represented at almost identical levels, irrespective of the catalyst or wavelength used. The mean composition is as follows, 21.5 ± 1.1% for group **I**, 28.5 ± 1.1% for group **II**, 50.0 ± 2.1% for group **III**, as stated earlier group **IV** does not contribute

significantly. A bar plot is provided showing the composition of each sample in Figure 5.10.

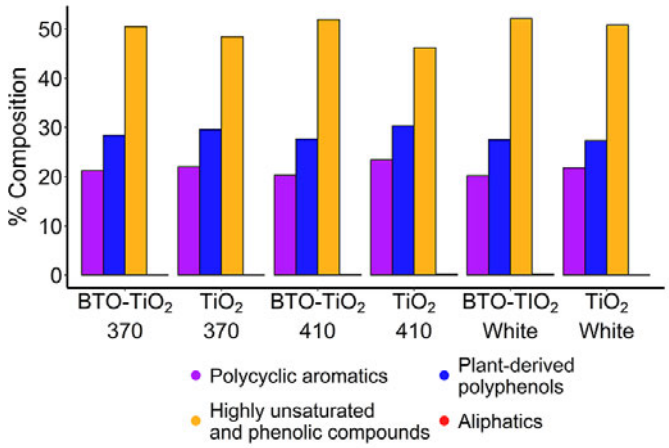


Figure 5.10: The percentage of molecular formulae removed by both catalysts at individual wavelengths for each compound class relative to the total number of removed formulae.

This consistent composition implies that the removal mechanism is identical regardless of the catalyst or wavelength and that the increased breakdown efficiency achieved by the BTO - TiO_2 catalyst and shorter wavelengths is due to a higher concentration of radicals produced or greater stability of the produced holes (h^+). Despite this, a question still remains, do the excess compounds removed at shorter wavelengths show different molecular characteristics from those removed by the longer wavelengths and similarly, those removed which are unique to each catalyst.

To answer these questions, the compounds removed by BTO - TiO_2 under different LEDs were inspected using UpSet plots (Figure 5.11a). The analysis was then repeated for TiO_2 (Figure 5.11b) and finally by considering both catalysts and all wavelengths together (Figure 5.11c). A quick inspection of the UpSet plots in Figures 5.11a and 5.11b reveals that only single digit numbers of formulae were uniquely removed by the 410 nm or white LED for both catalysts. In other words, the majority (average of 99%) of compounds that were removed, were also removed by the next shorter wavelength, regardless of which catalyst was used. For both catalysts, around 70% of the molecular formulae were removed at all wavelengths; the remaining 30% were cumulatively removed only at 370 nm or both 370 and 410 nm.

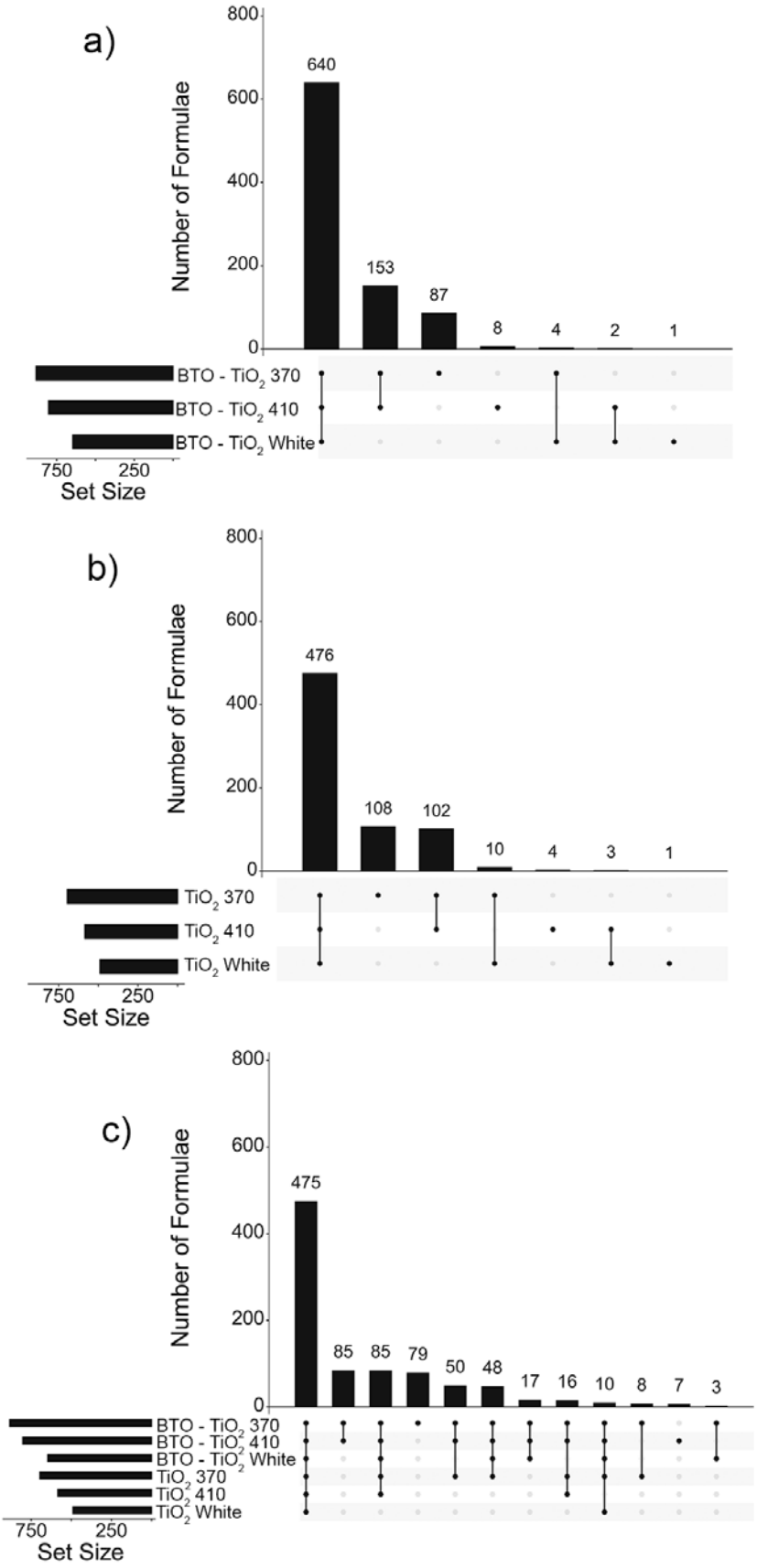


Figure 5.11: UpSet plots for the lost molecular formulae common to all time points in samples irradiated with three LEDs in the presence of (a) BTO - TiO_2 (b) TiO_2 and (c) the UpSet plot combining both groups of samples

A comparison of the molecules removed by BTO - TiO_2 under all LEDS (640 formulae) with those removed exclusively by 370 nm (87 formulae) or by both 370 and 410 nm (153 formulae) shows that the excess formulae removed at the shorter wavelengths are characterised by a lower DBE, lower oxygen class and a lower molecular weight than those removed by all three light sources, as seen in Figure 5.12. A similar trend was observed was observed for the pristine TiO_2 catalyst, see Figure C.12.

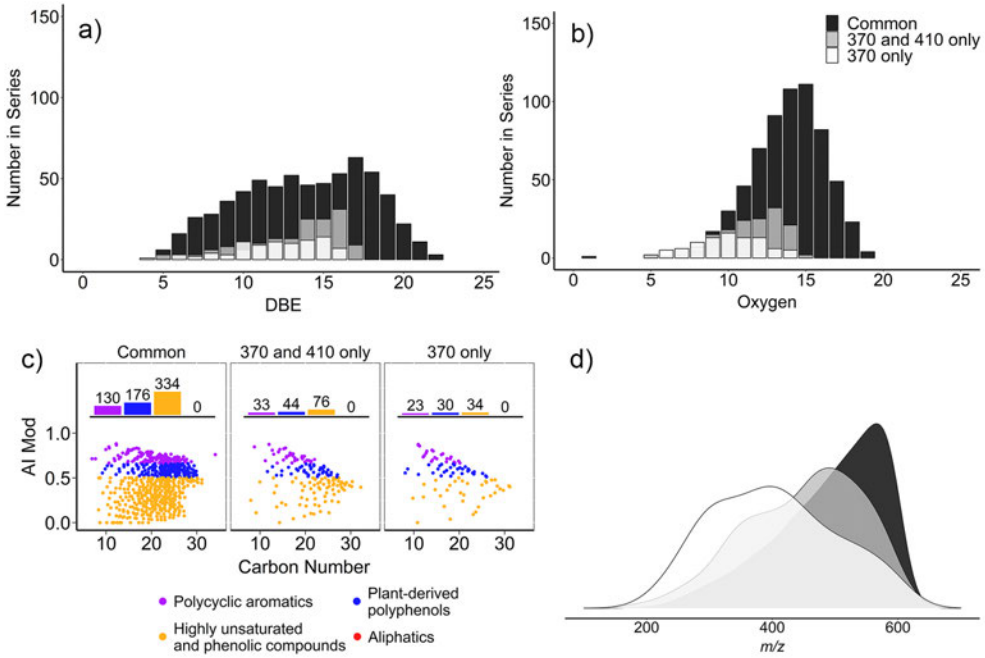


Figure 5.12: Molecular characteristics of lost formulae which are common to all three wavelengths (black, 640 formulae), unique to 370 and 410 nm, (gray, 153) or unique to 370 nm (white, 87), in the presence of the BTO - TiO_2 catalyst. a) a DBE plot, b) oxygen series, c) AI_{mod} plot with a bar plot of individual compound classes and d) m/z density plot normalised independently to the sum of all points in each group.

When focusing on the differences between the two catalysts, the UpSet plot in Figure 5.11c shows that no formulae were uniquely removed by the TiO_2 at any wavelength and that all formulae removed by this catalyst were also removed by the BTO - TiO_2 catalyst at some wavelength; on the other hand 193 formulae were uniquely removed by the BTO - TiO_2 catalyst at any wavelength. The molecular characteristics of these 193 are compared against the 694 formulae which are removed by both catalysts under any LED, below in Figure 5.13.

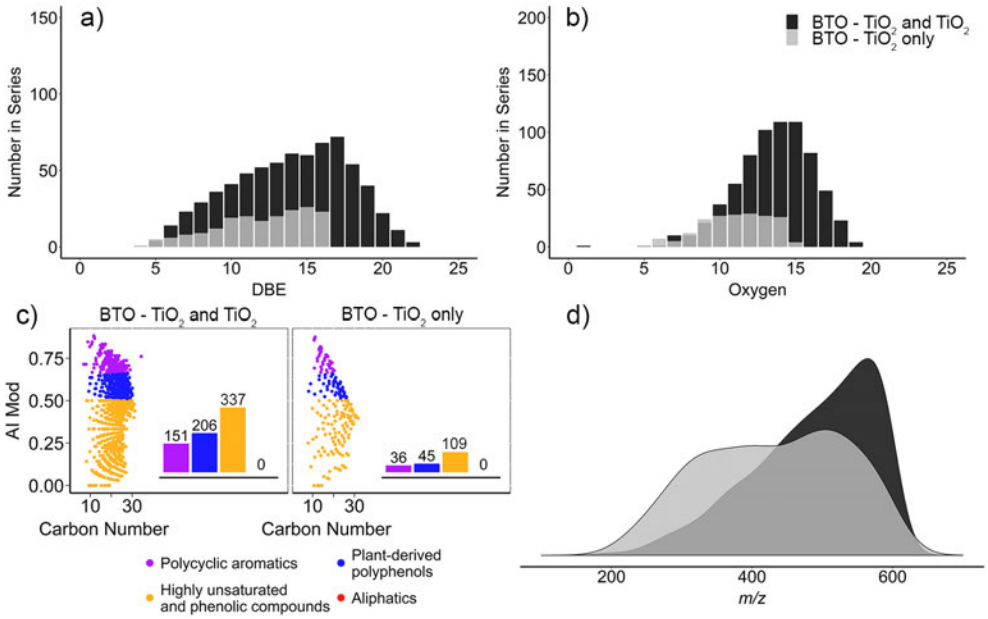


Figure 5.13: Molecular characteristics of formulae removed by any length of irradiation at some wavelength. A comparison of common molecular formulae removed by BTO - TiO_2 and TiO_2 (black, 694 formulae) and the BTO - TiO_2 only removed formulae (grey, 190 formulae). a) DBE; b) oxygen series; c) AI_{mod} plot with histogram of individual compound classes; d) molecular weight density, normalised independently to the sum of all points in each group.

The DBE and oxygen series plots in Figures 5.13a and b respectively, show that species removed exclusively by the BTO - TiO_2 catalyst have a lower DBE and oxygen number than those removed by TiO_2. They also contained a larger proportion of low molecular weight ($<$ 400 m/z) species.

5.4.5 *Degradation by-products (i.e. produced compounds*

The degradation intermediates can be characterised as covering the entire mass range analysed, with the greatest density occurring around M_W 300 g/mol and having low DBE, AI_{mod} and oxygen count. Their numbers across all wavelengths and both catalysts, are more comparable than the numbers of removed compounds (Figure 5.4). A total of 544 and 562 unique new molecular formulae, which are common to all time points, were identified in at least one wavelength for the BTO - TiO_2 and TiO_2 catalysts, respectively. Their molecular characteristics, including DBE, oxygen numbers, AI_{mod} and molecular weights are also very similar, as shown in Figure 5.14.

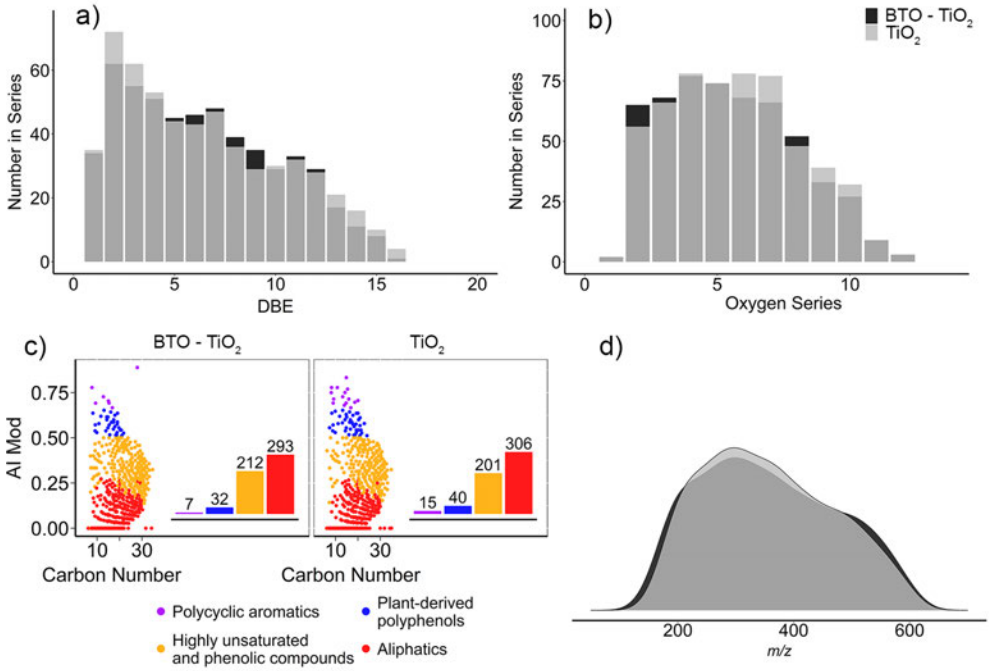


Figure 5.14: Comparison of newly identified molecular formulae between the BTO - TiO_2 (544) and TiO_2 (562) samples at any wavelength; a) DBE; b) oxygen series; c) AI_{mod} with histogram of individual compound classes, and d) molecular weight density normalised independently to the sum of all points in each group.

Nevertheless, for both catalysts, the number of new formulae depends on the irradiation wavelength, decreasing by 13% between the 370 nm and the white LED (based on the 3rd intersect in Figures 5.6 and C.2 - C.6). Interestingly, degradation intermediated differ between wavelengths. For the BTO - TiO_2 catalyst, only 332 formulae (equivalent to 61% of the new molecular formulae) are common to all three LEDs (Figure 5.16a). For the TiO_2 catalyst this number is even lower (262, or 47%), principally because 135 formulae are unique to the 370 nm LED (Figure 5.16b).

This diversity is also reflected in the representation of compound classes **II** - **IV** amongst the degradation intermediates. For the BTO - TiO_2 catalyst, 59% (370 nm LED) and 48% (White LED) were classed as aliphatic (**IV**), while the contributions from classes **II** and **III** increases accordingly (Figure 5.15). Class **I** compounds were not considered in this analysis as the numbers of produced polycyclic aromatics (**I**) were negligible. These observations highlight the fact that compounds are being degraded further under irradiation with the lower wavelengths and are more likely to form an aliphatic end molecule. The TiO_2 catalyst follows a similar trend (Figure 5.15).

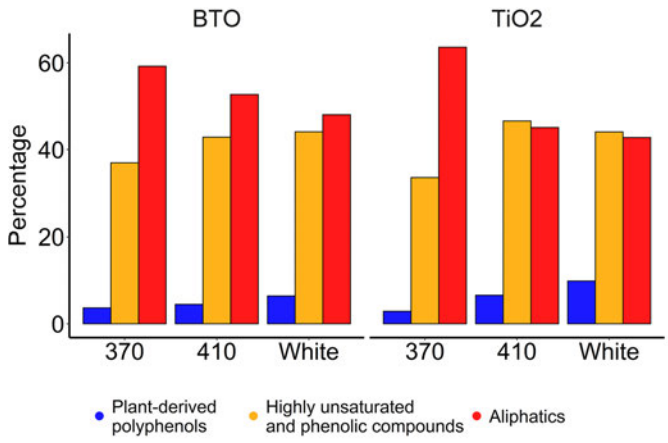


Figure 5.15: The composition of by-products produced by the BTO - TiO_2 catalyst at individual wavelengths per compound class relative.

However, the question still remains whether there are any differences in the intermediates produced by the two catalysts. An inspection of the UpSet plot combining the data for both catalysts showed that, out of 262 formulae common to all wavelengths in the TiO_2 catalyst samples (Figure 5.16b), 247 or 94% of these formulae are also shared by the BTO - TiO_2 catalyst (Figure 5.16c). Intermediates were identified that were only produced by one catalyst, e.g. 48 or 25 formulae at 370 nm for TiO_2 or BTO - TiO_2 respectively. Although these molecules are different, they have similar DBE and oxygen class distribution and the majority are classed as aliphatic molecules, 72% and 92% for BTO - TiO_2 and TiO_2 respectively. In case of BTO - TiO_2, the molecular mass distribution was slightly shifted to lower values compared to TiO_2. From a different perspective, out of the 135 formulae unique to TiO_2 under 370 nm (Figure 5.16b), only 48 remained unique to the TiO_2 and the 370 nm LED when BTO - TiO_2 was also considered (Figure 5.16c). The remaining 87 were found dispersed throughout intersections corresponding to the BTO - TiO_2 catalyst and one or more light sources. This observation indicates that a number of intermediates that require the most energetic light source in the presence of TiO_2 are also produced by BTO - TiO_2 at longer (less energetic) wavelengths.

Altogether, the intermediate molecules produced with different wavelengths or catalysts show greater heterogeneity than the degraded compounds.

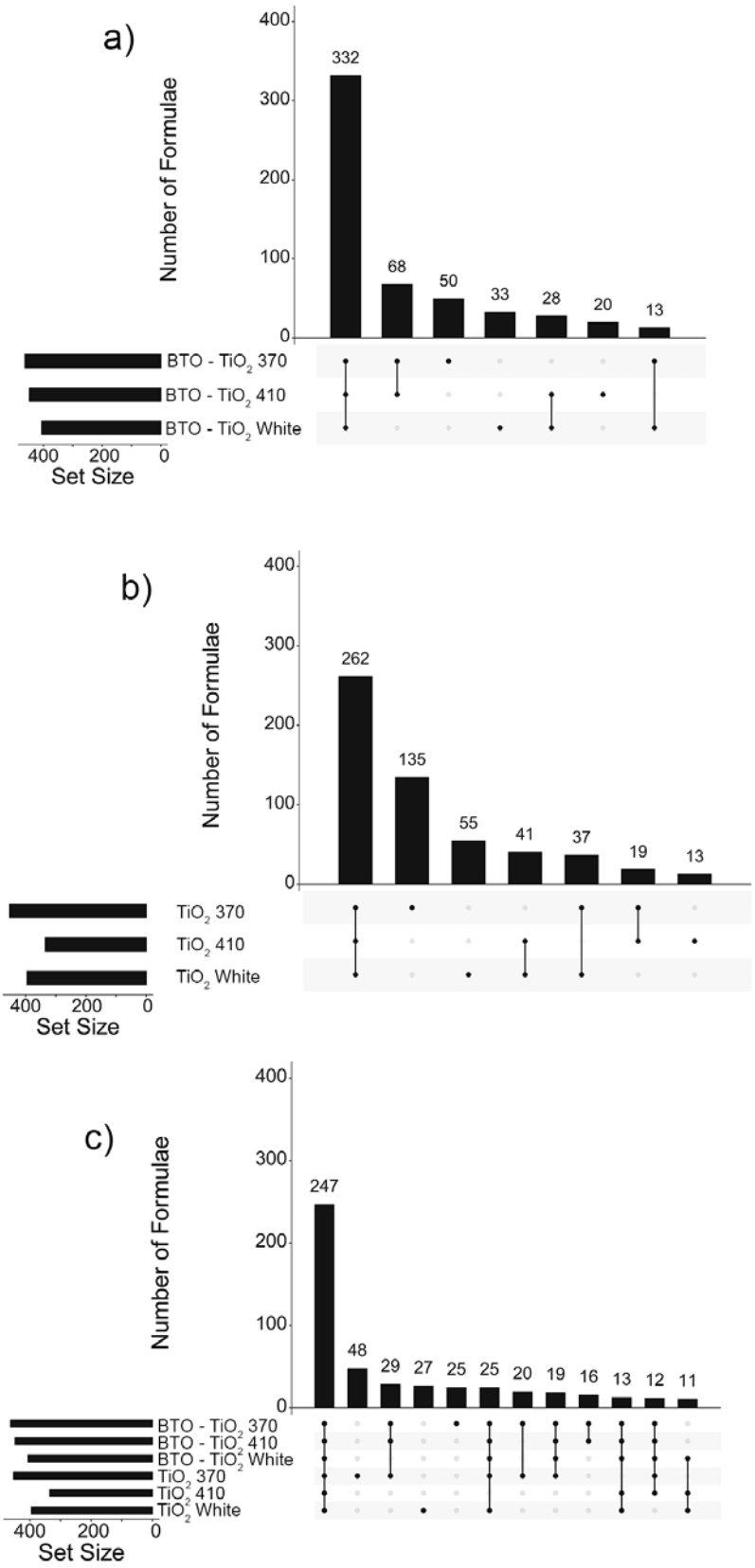


Figure 5.16: UpSet plots showing new molecular formulae common to all time points in samples irradiated with three LEDs in the presence of the a) BTO - TiO_2, b) TiO_2 catalysts and c) the UpSet plot combining the two.

5.4.6 *Adsorption of DOM to the BTO - TiO_2 catalyst and its relationship to photocatalysis*

Organic compounds are degraded by reactive hydroxyl radicals, which can be present as both free radicals present in the bulk solution ($\cdot OH_f$) and bound to the surface of the photocatalyst ($\cdot OH_s$),[132] as well as directly by holes generated on the catalysts surface. Adsorption of DOM molecules to the catalysts therefore plays an important role in the degradation process. A reduction in TOC at the beginning of a TiO_2-photocatalysed oxidation of DOM has been reported and attributed to the adsorption of DOM onto the surface of the catalyst.[133] Adsorption to BTO - TiO_2 catalysts is analysed here, aiming to determine molecular characteristics of the adsorbed species and to assess the role of adsorption in the photocatalysis.

A DOM sample was left in contact with the catalyst in the dark for 1 to 3 hours under conditions otherwise identical to the photocatalytic experiments. A loss of 129 molecular formulae was observed by the end of the first hour, gradually increasing and levelling off with the total of 180 and 212 formulae missing after 2 and 3 hours, respectively. The final time point represents 10.5% of the molecular formulae of the reference sample. The time course of the adsorption is compatible with a diffusion driven adsorption equilibrium on the surface of the catalyst. The majority of the adsorbed species had a high AI_{mod} index, high DBE, high oxygen and m/z (Figure 5.17), i.e. are mainly larger, oxygen rich aromatic species with characteristics identical to the compounds predominantly removed during photocatalysis. Out of 212 formulae, 56% belong to either compound class **I** or **II** and are thus easily degradable; 26 (or 12%) aliphatic compounds, which are unlikely to be degraded, were also adsorbed.

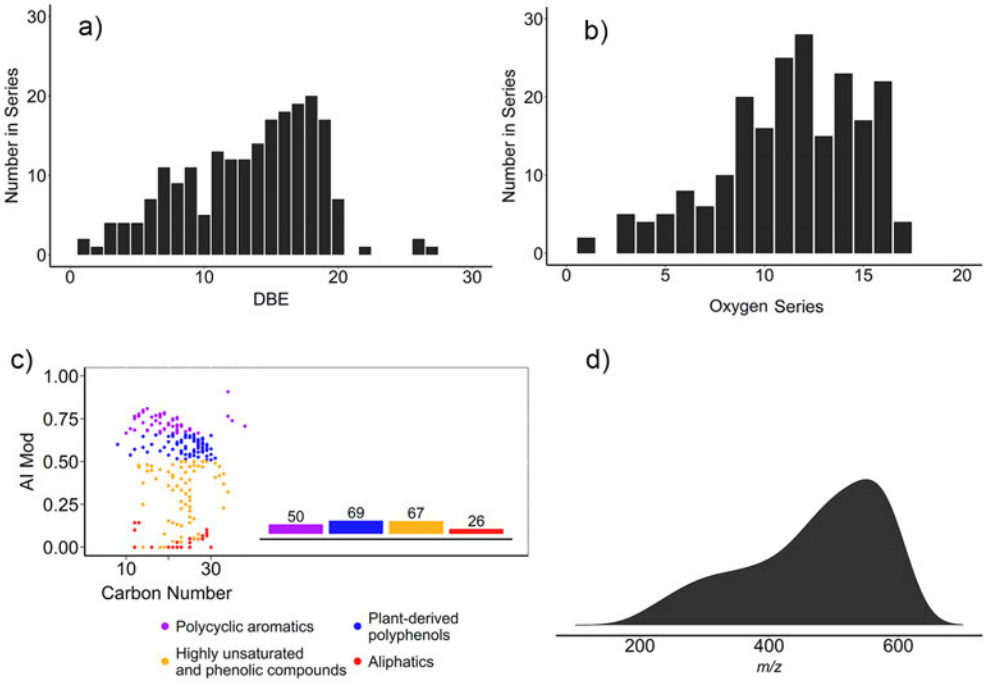


Figure 5.17: Molecular characteristics of 212 molecular formulae removed due to adsorption to the BTO - TiO_2 beads in the dark. a) DBE; b) oxygen series; c) AI_{mod} with a histogram of individual compound classes; d) a molecular weight density normalised independently to the sum of all points in each group.

The molecular formulae removed through the adsorption were then cross checked with the data from the degradation experiments. This comparison showed that after 1 hour of photocatalysis, approximately 25% of the molecular formulae missing due to adsorption were still detected; after 3 hours only 13% were present (Figure 5.18). The majority of these were aliphatic molecules, presumably displaced from the surface of the catalyst by compounds produced during photocatalysis. It can therefore be concluded that the high affinity, class **I - III**, molecules were degraded on the BTO - TiO_2 surface.

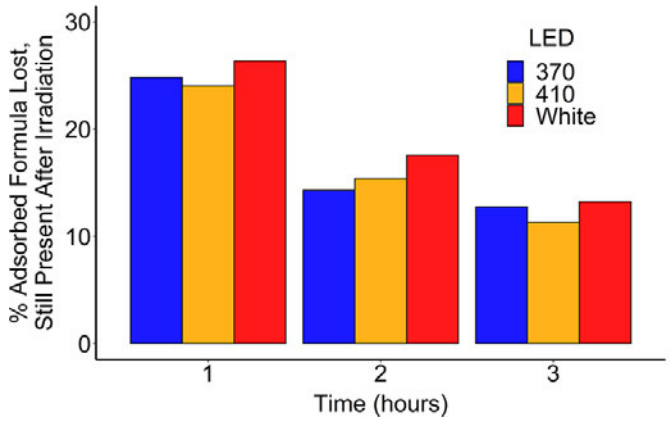


Figure 5.18: Percentage of formulae removed due to adsorption still present after irradiation at different wavelengths and times.

A more informed analysis is possible when the 212 formulae are looked at through the prism of compound classes. Molecules in classes **I**, **II** and **IV** showed similar affinity for BTO - TiO_2 with approximately 21% of formulae of the starting material removed through adsorption, while only 5% of class **III** molecules - highly unsaturated molecules and phenolics - were adsorbed. Does this 4-fold decrease in adsorption efficiency, of the class **III** molecules, lead to a similarly lower photocatalytic degradation efficiency?

Inspection of Figure 5.19, which shows (irrespective of wavelength) a degradation efficiency for class **III** molecules of only approximately 2 times lower than that of classes **I** and **II**. A possible explanation for the degradation efficiency of class III molecules being higher than expected, based on the adsorption data, is that in addition to degradation on the surface, molecules are also degraded by the radicals in the bulk solution.

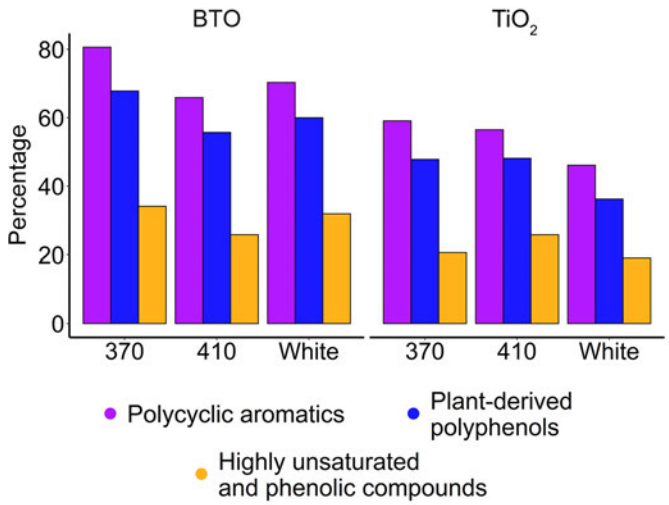


Figure 5.19: The relative composition of by-products produced by both the BTO - TiO_2 and the TiO_2 catalyst at individual wavelengths per compound class.

To summarise the interplay between adsorption and photocatalytic degradation, compounds with the highest affinity for the catalyst will bind first and, if susceptible to a radical attack, will be removed quickly. This leaves the lower affinity compounds to establish their adsorption equilibria and again be potentially degraded. Driven by Le Chateliers principle, this process will continue at a slower pace and degradation intermediates are created that further undergo transformations in a continuous manner. It is likely that photocatalytic degradation also takes place in the bulk solution.

5.4.7 *NMR analysis*

Figure 5.20 shows an overlay of two example 1D ^{1}H NMR spectra, clearly highlighting a drastic reduction in signal intensity across the entire spectral range, for the sample which has been irradiated (blue), with the majority of the remaining signals being present in more aliphatic regions, this is in comparison to the non-irradiated sample (purple).

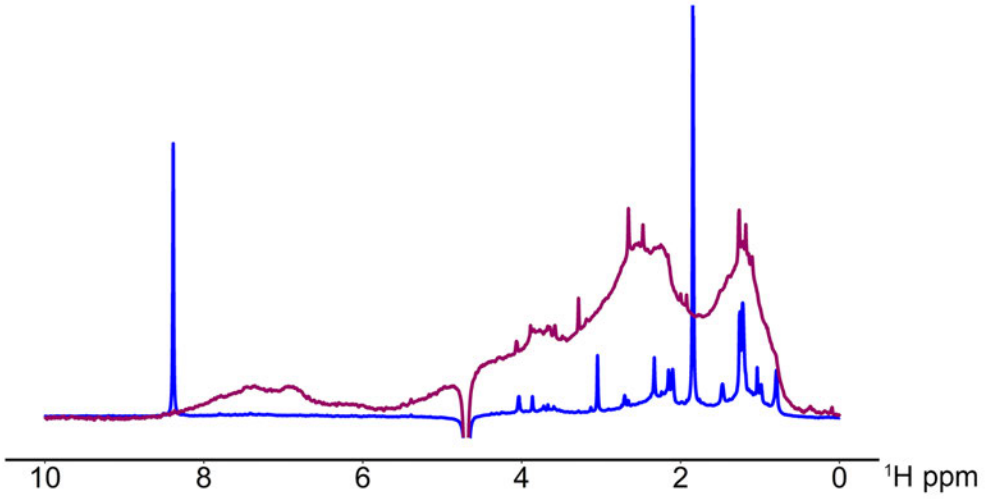


Figure 5.20: Superimposed 1D ^{1}H NMR spectra of the starting material (purple) and BTO - TiO_2 sample (blue) irradiated for 3 hours with the 370 nm LED.

Initial interpretation of ^{1}H NMR spectra of samples, irradiated by the three LEDs in the absence and presence of the catalysts, was done using Principle Component Analysis (PCA). Over 90% of the data could be explained with only 6 components, with PC1 explaining 67.2% of the data and PC2, 9.5% (Figure C.13). The PC1/PC2 scores plot (5.21a - c) are presented by highlighting samples from one light source at the time; PC1 loadings plot is shown in Figure 5.22.

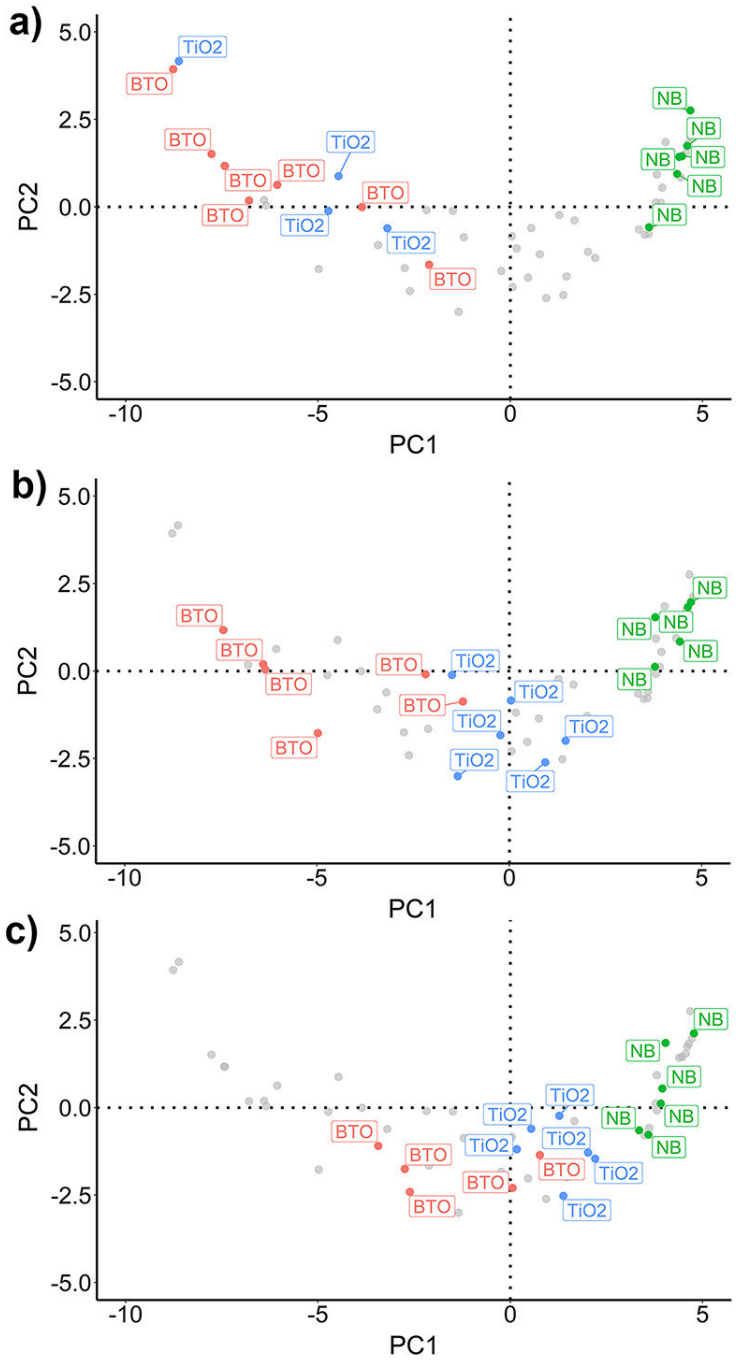


Figure 5.21: PCA scores of ^{1}H NMR data for no catalyst (NB, green), BTO - TiO_2 (BTO, orange) and TiO_2 catalyst (TiO2, blue). PC1 represents 67.2% of the variation within the dataset and PC2 represents 9.5%, giving a combined variance of 76.7%. a), b) and c) are the same scores plot but with with highlighted samples irradiated at 370 nm, 410 nm and the white LED respectively.

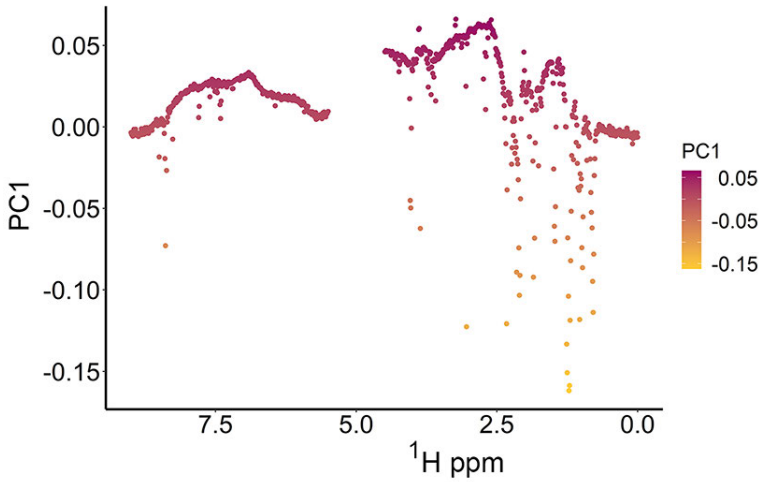


Figure 5.22: PC1 loadings plot of ^{1}H NMR PCA data.

For samples irradiated at 370 nm (Figure 5.21a), there is a clear separation between the no catalyst samples (NB, no beads samples) and those where the respective catalysts were present. This is to be expected as excitation of both the TiO_2 and the BTO - TiO_2 readily takes place at this wavelength, causing similar degradation, as indicated by previous MS analysis. However, irradiation at 410 nm (Figure 5.21b) caused the TiO_2 and the BTO - TiO_2 samples to form separate clusters, i.e. there are differences between the two catalysts, with the TiO_2 samples being positioned closer to the NB samples along PC1. This result was again to be expected as the heterojunction present in the BTO - TiO_2 catalyst, creates a heterojunction which is capable of excitation at lower energy wavelengths when compared to the pristine TiO_2. This remains evident still when looking at those samples irradiated with the white LED, (Figure 5.21c), the TiO_2 and the BTO - TiO_2 sample clusters both move closer to the NB samples, but the BTO - TiO_2 samples remain the furthest, indicating that there is greater change still when compared to the TiO_2 samples.

The PC1 loadings plot, (Figure 5.22) highlights which regions of the NMR spectra have contributed the most to the sample separation. The aromatic and carbohydrate (6.5 - 8.5 ppm and 4 - 5 ppm respectively) show positive values associated with these high abundance of these compounds in the NB samples. On the other hand, some aliphatic signals in the 0.5 - 2.5 ppm region have strong negative loadings, indicating their relative increase in the photodegraded samples. These and some other signals were identified with the help of a 2D ^{1}H - ^{1}H TOCSY and diffusion-edited NMR experiments (Figure 5.23) as belonging to long-chain fatty acids.

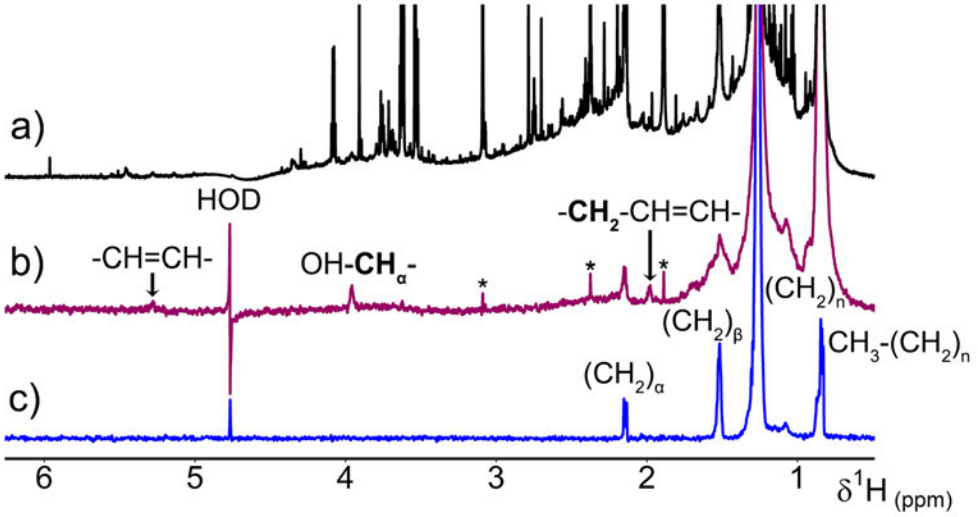


Figure 5.23: Identification of fatty acid chains in the BTO - TiO_2 sample irradiated at 370 nm for 3 hours. a) Partial 800 MHz 1D ^{1}H spectrum; b) diffusion-edited spectrum with signals of smaller molecules removed due to their faster diffusion (residual signals of small molecules are labelled with an asterisk); c) 1D trace through the $(CH_2)_n$ signals from a 2D ^{1}H - ^{1}H TOCSY spectrum.

These include typical signals of saturated (Figure 5.23c) unsaturated (5.29 and 1.98 ppm, Figure 5.23b) or hydroxylated (3.96 ppm, Figure 5.23b) fatty acids. Examples of fatty acids which are large and sufficiently soluble in water to be observed in diffusion-edited NMR spectra include, Myristic acid ($C_{14}H_{28}O_2$, M_w = 228 g/mol), 2-hydroxymyristic acid ($C_{14}H_{28}O_3$, M_w = 244 g/mol) or unsaturated ricinelaidic acid ($C_{18}H_{34}O_3$, M_w = 298 g/mol). Their molecular formulae were assigned in the FT-ICR spectrum of a respective sample and identified amongst the newly gained aliphatic compounds (class **IV**). The fact that these species are also observed by NMR shows that their appearance is genuine and not a result of a changed outcome of the competitive ionisation process at a particular stage of the photocatalytic degradation.

This is the first time that long-chain aliphatic fatty acids were identified amongst the products of the photocatalytic degradation of DOM. Their origin can be tentatively linked to the most depleted compound classes of polycyclic aromatics (**I**) and plant-derived polyphenols (**II**). Their transformations to fatty acids requires aromatic ring opening, double bond saturation and oxidation/ dehydroxylation - reaction mechanisms observed previously in the studies of photocatalytic degradation of single molecules.[134–136]

Turning our attention to the differences between the ^{1}H spectra: when the BTO - TiO_2 catalyst was used, intense signals of fatty acids were detected for all LEDs with the exception of the first time point of the irradiation with the

white light (Figure 5.24). With the TiO_2 catalyst, similar signal intensities were only observed for the 370 nm LED, while the other two LEDs produced weaker signals, showing very gradual growth between hours 1 and 3 of the experiment and never reaching the intensity seen in the 370 nm LED set up. NMR analysis thus demonstrated a superior performance of the BTO - TiO_2 catalyst and identified major degradation products.

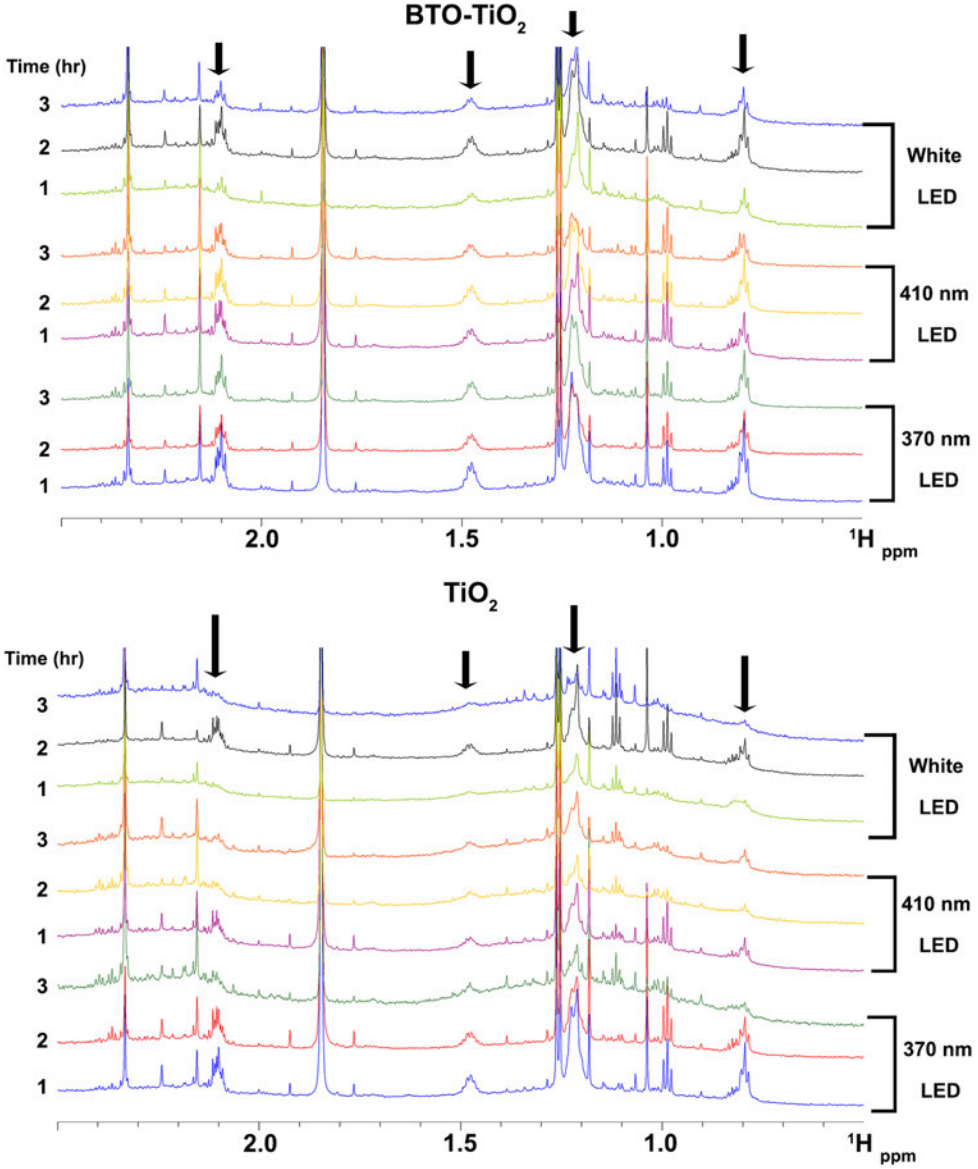


Figure 5.24: Partial 800 MHz 1D ^1H spectra over the course of the photocatalytic degradation experiment. Signals of fatty acid chains are indicated by black arrows.

5.5 CONCLUSIONS

Analysis of the MS data showed that, over the course of the three hours of irradiation, the number of molecular formulae assigned reduced by 15 - 25% for the BTO - TiO_2 catalyst, whilst this was less for the TiO_2 catalyst, between 0 and 16%. These numbers reflect the fact that, although a large number of molecules were removed, a number of new compounds were also produced. For the BTO - TiO_2 catalyst, between 0.52 and 0.63 new molecular formulae were identified for every molecular formula removed. For the TiO_2 catalyst the analogous ratios were between 0.65 and 0.80 for the 370 nm, 410 nm and white light sources.

The BTO - TiO_2 catalyst is able to utilise more wavelengths of light, thus reducing the demand for high energy UV lamps. Its performance, as reflected by the altered DOM composition, using the white light source was comparable to that of the TiO_2 catalyst at 370 nm. Inspection of the characteristics of individual molecular classes, showed that the removed compounds had high DBE, oxygen classes, aromaticity and were generally larger molecules. The opposite is true for the newly created compounds that were classified as smaller aliphatic molecules with fewer oxygens and double bonds. A high degree of similarity within the removed and gained groups of molecules, regardless of catalyst or light source, indicates that the photocatalytic mechanism is identical for both catalysts and all wavelengths. The shorter wavelengths and use of the BTO - TiO_2 catalyst leads to more efficient mineralisation of DOM. It is reassuring that NMR, as a quantitative technique reporting on much higher concentrations of compounds is able to confirm the results of the more sensitive, but qualitative technique, FT-ICR-MS.

Part IV

NMR STUDY OF CHLORAMINATION AND DEVELOPMENT OF ^{19}F NMR METHODS

CHLORAMINATION INVESTIGATION NMR SEQUENCE DEVELOPMENT

6.1 DECLARATION

The pulse sequences described here were designed and developed by myself for the application to fluorinated complex mixtures. The reduced dimensionality experiment however, was developed in unison with Richard York, a student in the Prof. Uhrin group and is included here for completeness. The chloramination experiments were designed and performed by myself.

6.2 OVERVIEW

In modern potable water supply/treatment, disinfection is arguably one of the most critical stages of treatment in terms of consumer safety, it also happens to be a process where the presence and the composition of DOM can have a significant adverse effect. Disinfection can be achieved through physical methods such as the application heat or the use of membranes with an extremely small pore size. These methods however are not feasible for the production of the substantial volumes required. Heat would be extremely costly, and although very fine membranes such as reverse osmosis membranes are available, only relatively low volumes can be produced by these methods. Even then, they would require significant investment and both methods would require some post treatment to prevent regrowth of organisms.

For this reason, the most common disinfection technique is the use of cheap effective chemical disinfectants such as chlorine, capable of not only disinfecting the water at point of treatment, but its residual often remains in the supply system, limiting regrowth. However, if DOM is present at the point of disinfection, it has the potential to react with the chosen disinfectant resulting in the formation of disinfection by-products (DBPs). The extremely heterogeneous nature of DOM is likely to produce a vast variety of potentially harmful by-products. Indeed some of the known DBPs have been shown to possess carcinogenic properties which has sparked significant academic and commercial interest in this area.

Chlorine is still the most widely used disinfectant globally, producing a variety of chlorinated by-products, some of which are regulated (THMs, HAAs). In order to avoid breaching these regulations, suppliers have been investigating alternative means of disinfection, one of which is chloramination. Chloramination involves the addition of both chlorine and ammonia, which results in the formation of monochloramine, a long lasting disinfectant that produces less of the regulated DBPs. However, with the addition of ammonia, there is an even greater variety of possible by-products that could be formed. Many of the nitrogeneous DBPs formed via chloramination have been shown to be even more genotoxic, cytotoxic or carcinogenic than many carbonaceous by-products.[137] Although, these compounds are produced in lower concentrations than their chlorination equivalents.

The majority of researchers studying DBPs do so using mass spectrometry, usually alongside gas chromatography (GC-MS). This affords the user with a mass and if acquired with a high enough resolution, a molecular formula. The retention time provided by the chromatography can also be used for identification, although this requires that the target compound be known and a standard is available to acquire the retention time for a respective compound. Fragmentation can also be applied, potentially leading to structural determination. This however, is not trivial and the fragments are usually not enough to confirm molecular structure. NMR is better suited although it struggles with mixtures. An approach to overcome this limitation is the main focus of this chapter.

When molecules are in isolation, the task of structural identification by NMR is rather straightforward, however, when mixed together, it quickly becomes far from trivial. With the increasing complexity of mixtures, the amount of signal overlap grows rapidly, preventing the identification of interacting nuclei, the essence of the NMR structural determination process.

Beyond decreasing the signal overlap by increasing dimensionality of NMR experiments, several other approaches can be exploited to separate signals and have been reviewed previously.[86,138] These include limiting the footprint of individual signals by removing their splitting caused by J interaction, selective signal excitation, or utilizing differences in the spin-spin relaxation or molecular diffusion to "edit" the spectra.

Another approach, recently pioneered in the analysis of very complex mixtures of soil organic matter relies on the introduction of isotopic labels to a subset of molecules.[139] As these tags are fully NMR active and modify only a subset of molecules, their signals stand out from the rest allowing them to report on

the chemical shifts of nuclei in their vicinity. The outcome is the structural characterisation of molecules that carry the tags.

When studying chemical processes that involve and/or produce mixtures of molecules, the possibility of working with tagged molecules is very promising. Here the tagging can take place before or after a chemical reaction took place. The former option is particularly attractive as it allows for a greater flexibility in choosing the tag. With this regard, fluorinated molecules stand out amongst other possible approaches mainly due to NMR properties of ^{19}F. High sensitivity, 100% natural abundance, large chemical shift dispersion and large and far-reaching spin-spin interactions make fluorination ideal for this purpose.

6.3 CHAPTER AIMS

This chapter has the following aims:

1) To develop a protocol for chloramination of DOM using NaOCl and NH_3Cl.

2) To apply this method to a model DOM sample.

3) To develop NMR experiments based around the ^{19}F nucleus, designed to interrogate the complex mixtures resulting from the chloramination reaction.

4) To determine the structure of compounds produced by chloramination of a model organic compound containing a single ^{19}F site.

6.3.1 *Experimental Methods*

The monochloramine stock solution was prepared fresh by dissolving ammonium chloride, in this case ^{15}N labelled, in MS grade water. Sodium hypochlorite was slowly added to make a molar ratio of 0.8:1, chlorine:ammonium. Like in refs[140] and[141], a 3:5 ratio of SRFA (as C, SRFA is approx 52.34% carbon[126]) to monochloramine (as Cl_2) was used and immediately after chloramine addition the pH of the solution was lowered to 7.5 with the addition of hydrochloric acid. After 5 days the reaction was stopped by addition of excess sodium thiosulphate ($Na_2S_2O_3$). The pH of the solution was then lowered to 2 using hydrochloric acid and pumped through a PPL SPE column with a water aspirator. After loading, the column was washed with 20 mL of acidified water before being eluted with methanol. The eluent was taken to dryness via rotary evaporation.

The sample was then redissolved in partially deuterated methanol (CD_3OH, not CD_3OD) in order to prevent exchange of NH protons with the deuterated solvent.

6.4 RESULTS AND DISCUSSION

6.4.1 *Initial Chloramination*

In order to investigate the by-products produced by chloramination, a lab experiment was devised taking influence from two studies in particular[140,142]. The first, reacts the well known riverine DOM standard Suwannee River Fulvic Acid (SRFA) with ^{15}N labelled chloramine in an effort to describe the nitrogenous by-products using ^{15}N NMR. This study is relatively old, published in 1992, and as such the authors did not have access to modern instrumentation. They only recorded 1D ^{15}N INEPT spectra, which had poor resolution. The second study is much more recent and again reacts SRFA with chloramine with the intent on using FT-ICR-MS to characterise DBPs. The reaction procedure used in this chapter is based on that reported in the second study.

SRFA was reacted with ^{15}N labelled chloramine as described in the experimental section and several spectra were recorded which highlight the extreme complexity of the product (Figure 6.1). A 2D ^{1}H - ^{15}N HSQC spectrum contains hundreds of cross peaks highlighting the very significant nitrogen incorporation into the sample. Three distinct groups, labelled here from a-c, of resonances are visible, which based on the chemical shift values are assumed to be a) amines b) primary amides and c) secondary amides.

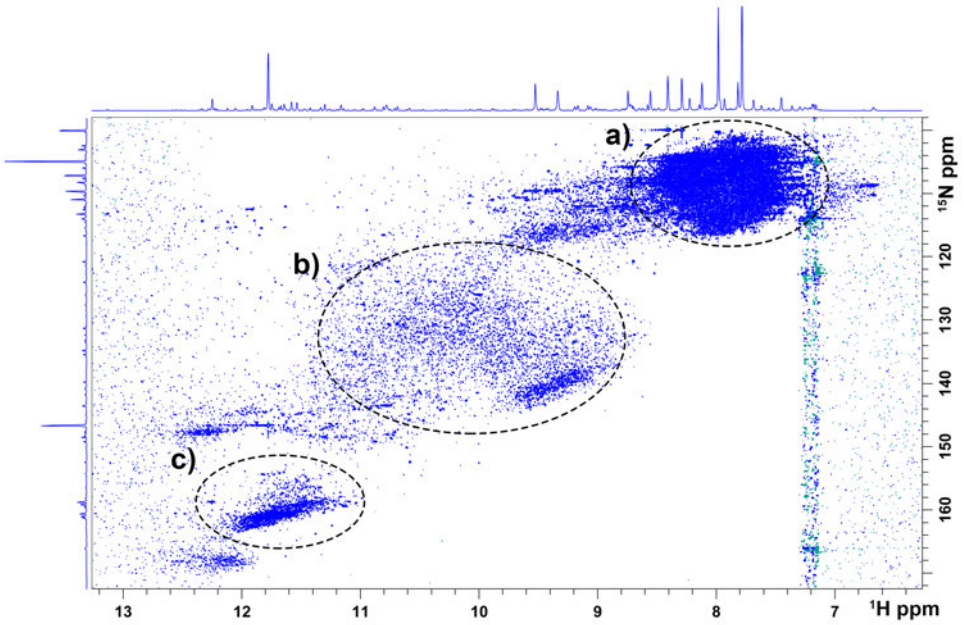


Figure 6.1: A ^{1}H - ^{15}N HSQC highlighting the extent of nitrogen incorporation into this complex mixture. regions highlighted are assumed to be **a)** amines **b)** primary amides and **c)** secondary amides.

Due to the extreme complexity of the sample, it was decided that either we needed to increase the resolution of the experiments through the use of multidimensional NMR experiments, or reduce the complexity of the system. We opted for the latter, and instead of chloraminating a complex mixture, which would give a correspondingly complex product, we decided to first chloraminate a pair of small organic acids considered to have representative structural features of molecules typically found in DOM. The acids chosen were gallic acid (**1**) and ferulic acid (**2**).

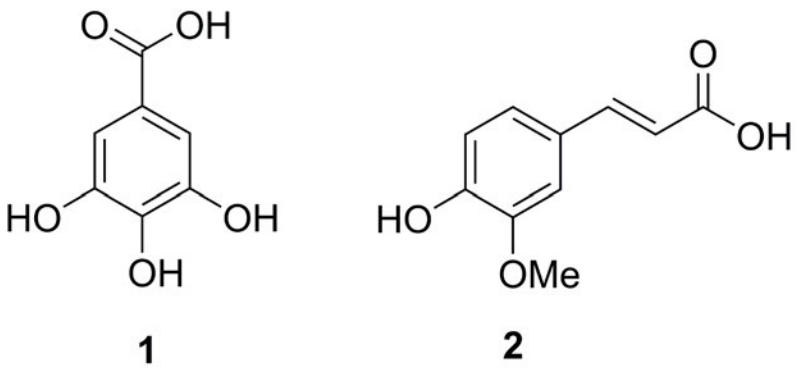


Figure 6.2: Structures of gallic acid (**1**) and ferulic acid (**2**).

Surprisingly, even when only two organic acids were chloraminated (individually) using the same protocol, the resulting ^{1}H - ^{15}N HSQC spectra looked extremely complicated (Figures 6.3a and b)

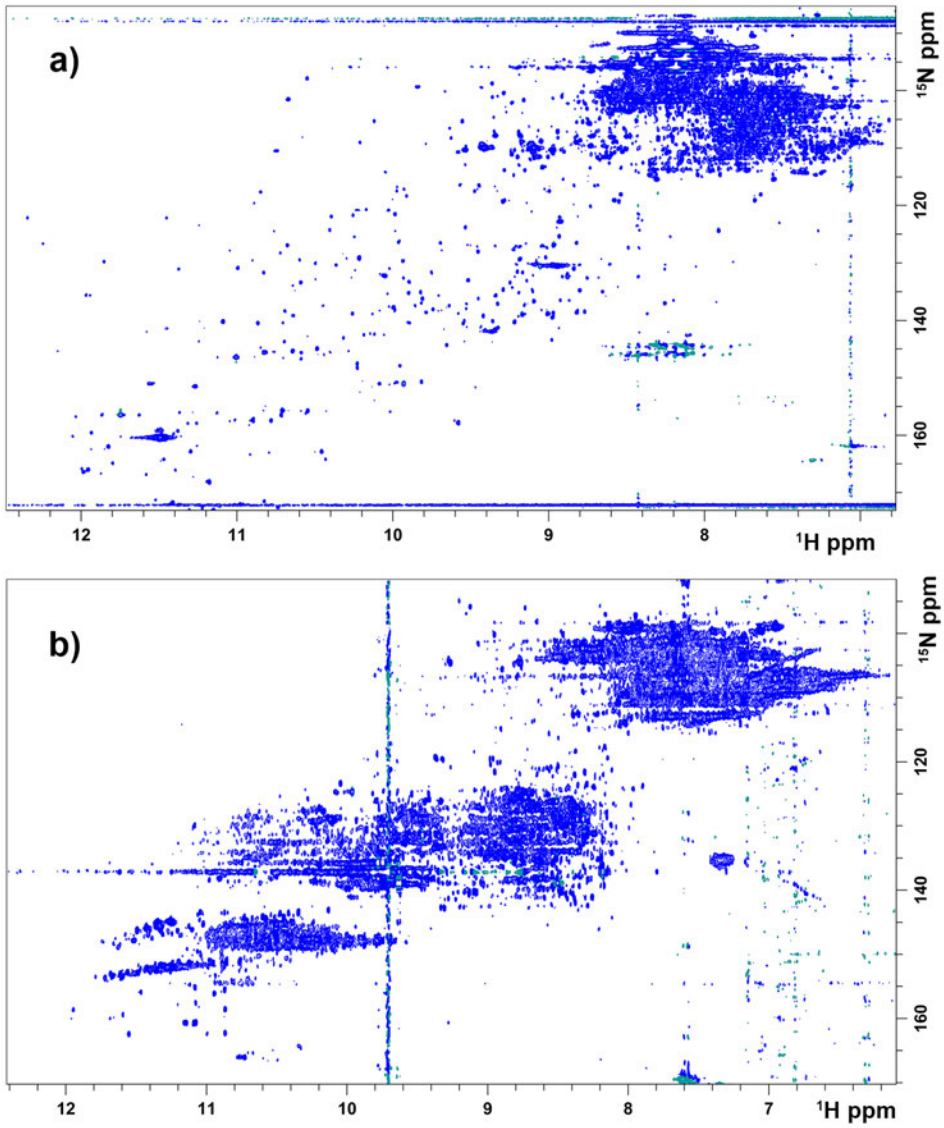


Figure 6.3: 2D ^{1}H - ^{15}N HSQC spectra of **a)** Gallic acid (**1**) and **b)** Ferulic acid (**2**).

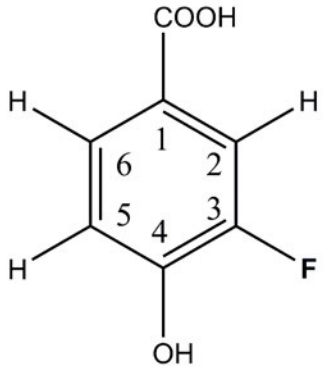
Figure 6.3 shows that less primary and secondary amides resulted from the chloramination of compound **1** compared to compound **2**. This may highlight that systems which are more conjugated can react with the disinfectant to a greater extent and are more likely to result in a greater variety of DBPs.

Figure 6.3 indicates that, again, we find ourselves in need of greater resolution to determine the structures of some of these by-products. Reducing the complexity

of the starting material unfortunately did not reduce the complexity of the resulting DBP mixture to the extent needed in order to truly interrogate the sample.

Another way of achieving greater resolution would be to incorporate an extra dimension to the NMR experiments, an experiment which is applicable to this situation would for example correlate ^{1}H, ^{15}N and ^{13}C. Such an experiment could start on protons, transfer magnetisation to nitrogen via $^{1}J_{NH}$, followed by the transfer via $^{n}J_{CH}$ to carbon and all the way back to protons for detection. The resulting spectrum can be displayed as a cube, with ^{1}H, ^{15}N and ^{13}C axes. Indeed, protein NMR spectroscopists use such an approach to correlate the chemical shifts of H, N and CO atoms of the protein backbone in a 3D HNCO experiment.[143] The main issue regarding higher dimensional experiments is that the indirectly detected nuclei need to be sampled individually, resulting in much longer experimental times. The spectra are also inherently more complex and therefore also take longer to analyse. In addition, unlike in fully labelled proteins, ^{13}C is present at natural abundance, lowering the sensitivity of such experiments on nitrogenated compounds.

In order to overcome the hindrance of spectral overlap, it was decided to incorporate ^{19}F into the starting material **3**, 3-Fluoro-4-hydroxybenzoic acid, a small organic acid similar to **1** and **2**. If the ^{19}F is not removed by the reaction with chloramine, or the compounds are not oligomerised, the resulting DBPs should only have one ^{19}F nucleus, meaning we do not have to worry about ^{19}F - ^{19}F couplings. If the reaction leads to oligomerisation, the DBP's will contain several ^{19}F nuclei, however they are likely to be too far apart to exhibit any coupling with each other.



3

Figure 6.4: Structure of the starting material (**3**), 3-Fluoro-4-hydroxybenzoic acid.

6.4.2 *Developing NMR experiments for mono-fluorinated compounds*

6.4.2.1 *2D ^{1}H - ^{19}F HETCOR and HMBC*

Following the acquisition of ^{1}H coupled and ^{1}H decoupled ^{19}F spectra of a mixture of monofluorinated compounds, establishing the ^{1}H - ^{19}F correlations is naturally considered as the next step. As discussed in,[144] not all spectrometers currently in use are capable of performing experiments which pulse both on ^{1}H and ^{19}F in a single experiment, nevertheless the situation is gradually improving.

If available, the choice is between three experiments: hetero - COSY, HETCOR or HMBC. All can be acquired by directly detecting either ^{1}H or ^{19}F , the latter is usually preferred as indirectly sampling the large ^{19}F chemical shift range in F1 can be very time intensive. All three of these approaches have advantages and disadvantages.

Hetero - COSY does not contain any fixed delays and produces pure phase spectra, but yields antiphase multiplets split in both the F1 and F2 dimensions by ^{1}H - ^{19}F couplings, these are further split by ^{1}H - ^{1}H couplings in F1 resulting in even more complex multiplets and/or broadening. Non-refocused HETCOR yields pure - phase antiphase signals in F2 (^{1}H - ^{19}F J) and the cross peaks are split/broadened by ^{1}H - ^{1}H couplings in F1. Their intensity depends on the value of the transfer function governed by both the active J_{HF} and passive J_{HH} coupling constants. Non - refocused HMBC has similar attributes to HETCOR, but its transfer function depends on the active and passive J_{HF} coupling and not the J_{HH} couplings. Both of these sequences contain 180° pulses, which could be limiting if only longer ^{19}F pulses (>15 μs 90° pulse) are available, this however is not an issue when broadband 180° pulses are used. As mentioned above, all three of these experiments can be performed with the direct detection of ^{19}F or ^{1}H , the former was used due to the need for more resolution when working with these complex mixture samples. However, ^{1}H - ^{1}H coupling constants are not easily obtained via ^{19}F detected experiments.

In order to maximise the sensitivity and applicability of these techniques to the large ^{19}F chemical shift range seen in these complex mixture samples, the HETCOR and HMBC experiments were performed in a phase sensitive mode using broadband inversion pulses. Sufficient for the majority of applications, 1 ms 180° CHIRP pulses[145] were used that invert a 110 kHz spectral window (i.e. ≈ 300 ppm for ^{19}F on a 400 MHz NMR spectrometer), see Figure 6.5.

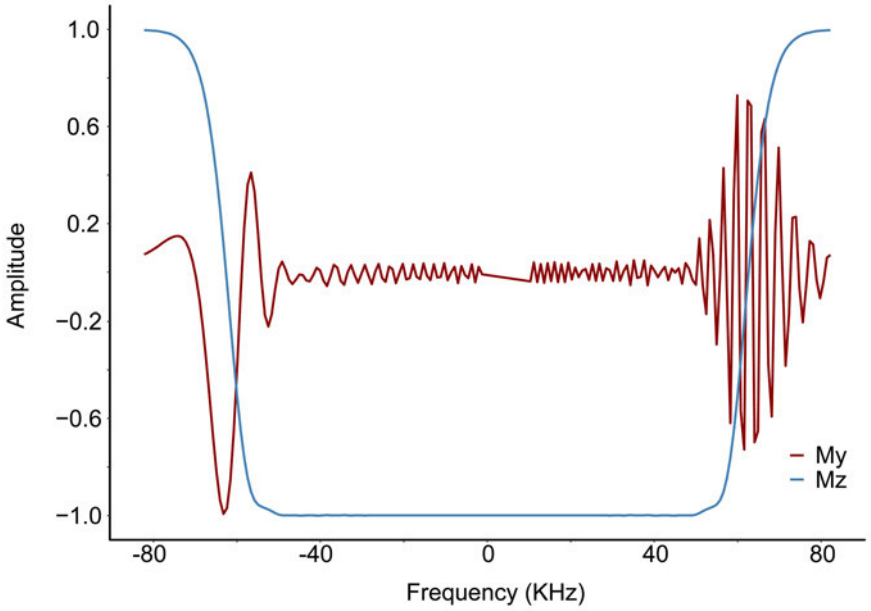


Figure 6.5: Excitation profile of 1 ms CHIRP inversion pulse.

As two such pulses were applied consecutively, the phase evolution of transverse magnetization during the pulses is eliminated, allowing their use not only for inversion (HETCOR) but also for refocusing (HMBC). Accurate timing ensured that any ^{19}F chemical shift evolution was refocused (HMBC) prior to acquisition, allowing pure phase spectra to be recorded. No attempt was made to refocus the magnetization prior to ^{19}F detection due to the large spread of the J_{HF} sizes. Many J_{HF} constants were determined from antiphase splitting of cross peaks in F2 from both HETCOR and HMBC spectra. The modified pulse sequences for these already well-established experiments are shown in Figure 6.6. The pathway of magnetisation for these experiments is given in Equation 6.1a- c.

$$\begin{aligned}
&\text{a) } {}^{1}\mathrm{H} \xrightarrow{t_1(^{1}\mathrm{H}),\,{}^{n}J_{HF},\,{}^{n}J_{HH}} {}^{19}\mathrm{F}(t_2) \\
&\text{b) } {}^{1}\mathrm{H} \xrightarrow{t_1(^{1}\mathrm{H}),\,{}^{n}J_{HH}} \xrightarrow{\Delta_1+\Delta_2;\ {}^{n}J_{HF}} {}^{19}\mathrm{F}(t_2) \\
&\text{c) } {}^{19}\mathrm{F} \xrightarrow{\Delta_1+\Delta_2;\ {}^{n}J_{HF}} \xrightarrow{t_1(^{1}\mathrm{H})} {}^{19}\mathrm{F}(t_2)
\end{aligned} \tag{6.1}$$

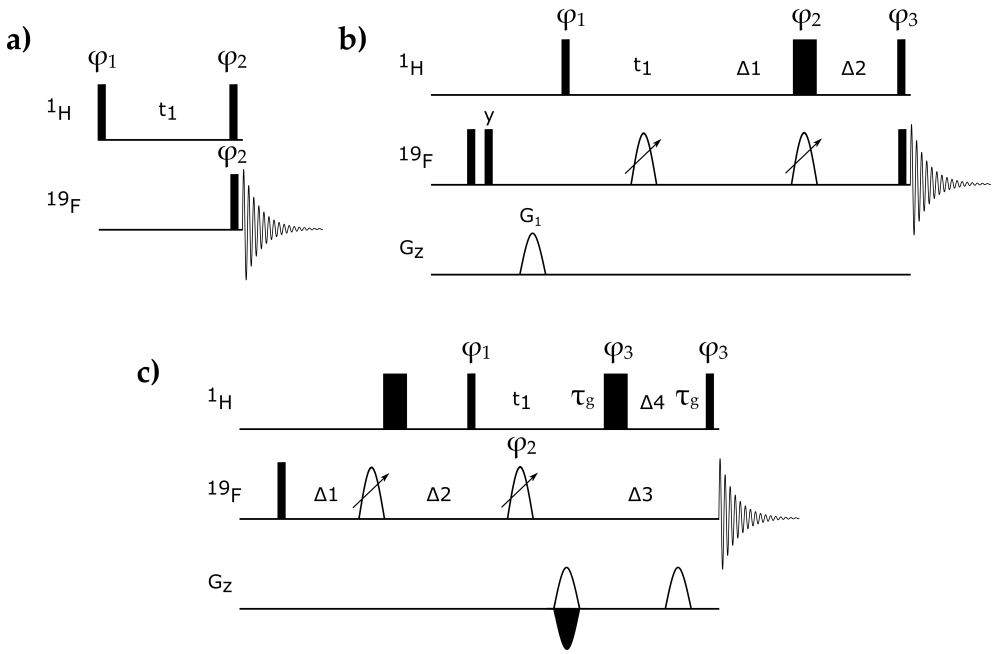


Figure 6.6: Pulse sequences of **a)** 2D ^{19}F - ^{1}H COSY, **b)** 2D ^{19}F - ^{1}H HETCOR and **c)** 2D ^{19}F - ^{1}H HMBC. The thin and thick filled rectangles represent high power 90° and 180° pulses respectively. The 1 ms adiabatic CHIRP pulses (p12) are indicated by an inclined arrow. Unless stated otherwise, pulses were applied from the x axis.

In **a)** the initial t_1 evolution delay was set to 6 μs. The following phases were used: φ_1 = x, 2(-x), x, y, 2(-y), y; φ_2 = 2(x, -x), 2(y, -y) and the receiver phase = x, 2(-x), x, y, 2(-y), y.

In **b)** the following delays were used: $\Delta_1 < \frac{1}{4} J_{HF}$ (optimised for 20 Hz in this case), $\Delta_2 = \Delta_1$ + p12 + $t_1(0)$, where $t_1(0)$ is the initial t_1 evolution delay, set to 6 μs. The following phases were used: φ_1 = 2x, 2(-x); φ_2 = 4x, 4(-x); φ_3 = y, -y and the receiver phase = x, 2(-x), x. Phase φ_1 was incremented in the STATES - TPPI manner between increments.

In **c)** the following delays were used: $\Delta_1 < \frac{1}{4} J_{HF}$ (optimised for 20 Hz in this case), $\Delta_2 = \Delta_1$ - p4 + $(2/\pi)$p1 + $2\tau_g$ + Δ_4 - p3; $\Delta_3 = 2\tau_g$ + p4 + Δ_4 and Δ_4 = p12 + $t_1(0)$, where p3 and p4 are the 90° and 180° ^{1}H pulses, p1 is the 90° ^{19}F pulse, τ_g is the PFG and recovery delay (1.2 ms) and $t_1(0)$ is the initial t_1 evolution delay, 6 μs. The following phases were used: φ_1 = 2x, 2(-x); φ_2 = x, -x; φ_3 = 4x, 4(-x) and the receiver phase = 2(x, -x) 2(-x, x). Echo-anti-echo protocol was used for sign discrimination in F_1 by changing the polarity of the PFG's.

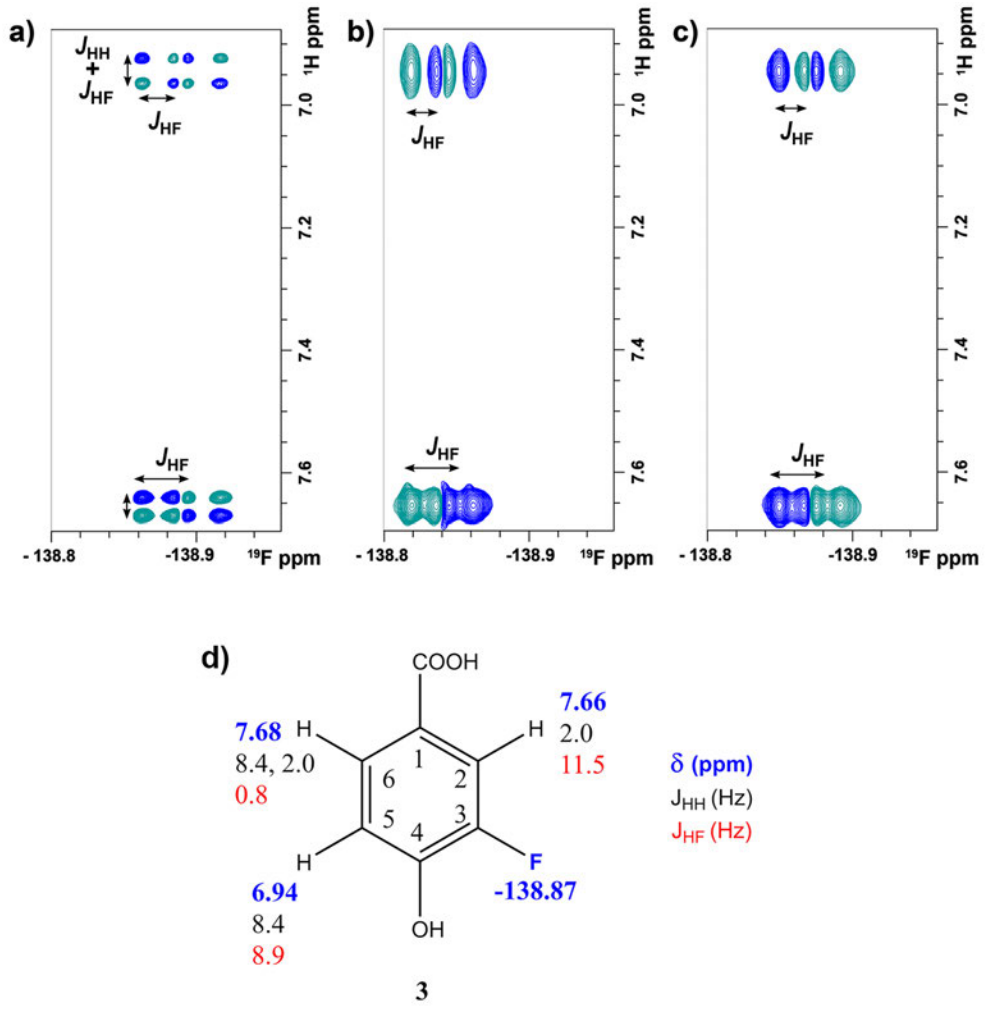


Figure 6.7: Three spectra of the starting material (**3**), **a)** ^{19}F - ^{1}H Hetero-COSY **b)** ^{19}F - ^{1}H HETCOR and **c)** ^{19}F - ^{1}H HMBC. **d)** The structure of the starting material (**3**) showing the ^{1}H / ^{19}F **chemical shifts (δ, ppm)**, the ^{1}H - ^{1}H coupling constants and the ^{1}H - ^{19}F coupling constants. Note proton 6 information was obtained from 1D ^{1}H experiments due to overlap with proton 2 in the HETCOR experiment.

As can be seen by Figure 6.7a, the Hetero - COSY cross peaks present themselves as complex multiplets, due to being split in F1 by both the J_{HF} and J_{HH} couplings and further split in F2 by the J_{HF}. By collapsing some of these couplings signal to noise is improved, which makes the spectra easier to interpret. Because of this, only the HETCOR and HMBC variations were implemented further. Spectra shown in Figures 6.7b and c, provide ^{1}H and ^{19}F chemical shifts and by analysing the cross peak splitting patterns, the J_{HF} constants. The antiphase splitting corresponds to the active coupling and the in-phase splitting corresponds to the passive coupling. Information obtained for the starting material (**3**) is highlighted in Figure 6.7d.

It should be noted that the transfer of magnetisation between nuclei is dependant on the size of the coupling constants, the HETCOR and HMBC experiments above are optimised for a 20 Hz coupling, because of this there was poor transfer to proton 6 and very low signal. The 0.8 Hz from the fluorine to proton 6 was obtained through comparisons of proton experiments with and without fluorine decoupling.

6.4.2.2 *2D ^{1}H - ^{19}F TOCSY-HETCOR*

Protons are only detected using this experiment if they have a direct, sizeable coupling with the fluorine. This however is not always going to be the case, especially if oligomerisation occurs, therefore, a 2D ^{1}H - ^{19}F TOCSY-HETCOR experiment was designed. This experiment allows the detection of protons, which do not directly couple to the fluorine, as long as they are within a spin-system with a proton which does.

The chemical shifts of all the protons are labelled before their magnetisation is spread through the networks of coupled spins by a DIPSI-2 spin-lock. Part of the magnetisation that reached fluorine-coupled protons is then transferred to the fluorine for detection. Comparing the resulting spectrum to that of the 2D ^{1}H - ^{19}F HETCOR spectrum reveals any additional protons belonging to the same spin system.

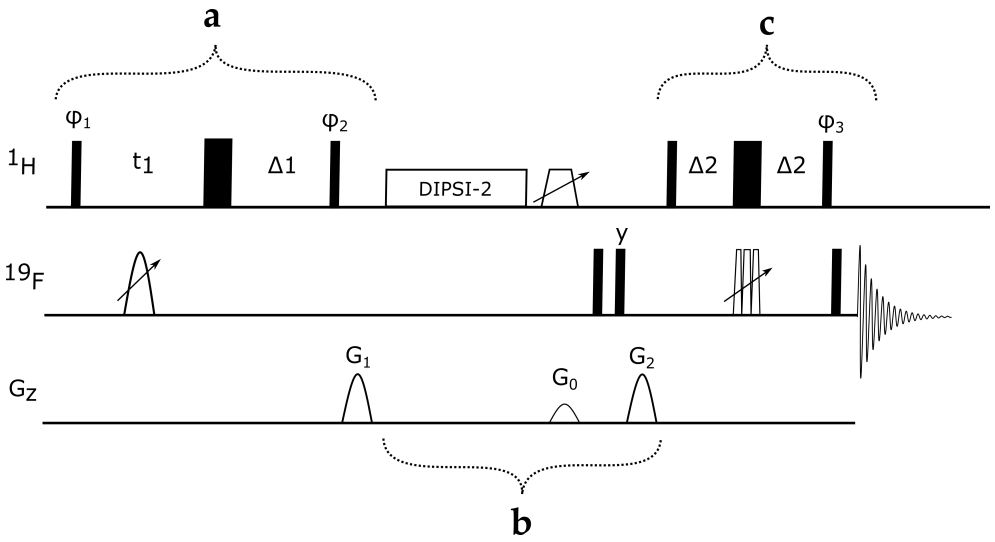


Figure 6.8: Pulse sequence of 2D ^{1}H - ^{19}F TOCSY-HETCOR experiment. Thin and thick filled rectangles represent high power 90° and 180° pulses respectively. The sine shaped pulse with an inclined arrow represents a 180° 1 ms adiabatic pulses, the trapezoid with an inclined arrow represents a 2 ms composite 180° pulse, the low level trapezoid with an inclined arrow represents a 20 ms adiabatic inversion pulse. $\Delta_1 = \mathrm{pw}_{180}(^{19}\mathrm{F}) + 2\mathrm{t}_1(0)$, where $\mathrm{pw}_{180}(^{19}\mathrm{F})$ is the ^{19}F 180° inversion pulse and $\mathrm{t}_1(0)$ is the initial t_1 evolution delay (4 μs). $\Delta_2 = \frac{1}{4}\ J_{HF}$ (optimised for 20 Hz in this case). The following phases were used: φ_1 = 2x, 2(-x); φ_2 = 4x, 4(-x); φ_3 = y, -y and the receiver phase = x, 2(-x), x , -x, 2(x), -x.

The below equation describes the flow of "meaningful" magnetisation throughout the experiment.

$$^{1}\mathrm{H} \xrightarrow{t_1(^{1}\mathrm{H}),\ ^{n}J_{HH}} \xrightarrow[\mathrm{DIPSI2}]{^{n}J_{HH}} \xrightarrow{^{n}J_{HF}}\ ^{19}\mathrm{F}(t_2) \tag{6.2}$$

In part **a** of the experiment shown in Figure 6.8a, the initial 90° pulse rotates the proton magnetisation to the y plane (e.g. y, although this alternates with phase cycling). Their chemical shift is then labelled during a variable t_1 period, at the centre of which is a 180° fluorine pulse, which prevents the evolution of any J_{HF}. Δ_1 is equal to t_1 + pw_{180} creating a spin echo, returning the magnetisation back to where it was after the initial 90° pulse (y). Another 90° proton pulse puts the magnetisation into z and is followed by a purging gradient to de-phase any remaining $x - y$ magnetisation.

A DIPSI-2 sequence (Figure 6.8 part **b**) results in the spread of magnetisation amongst coupled protons. A z-filter is implemented to defocus any multiple quantum (MQ) ^{1}H coherences that may have evolved. The z filter also incorporates a suppression of the natural ^{19}F magnetisation as described for the HETCOR experiment above. This minimises the cancellation artefacts in the experiment that relies on phase cycling to select magnetisation that originated in protons.

In part **c**, a 90° proton pulse returns the magnetisation into y, two Δ_2 delays with a combined length of $\frac{1}{2}$J results in the development of ^{1}H - ^{19}F antiphase magnetisation which is then transferred to the fluorine by two 90° pulses, one on proton the other on fluorine. There are simultaneous 180° pulses on proton and fluorine after the first Δ_2 delay in order to prevent the evolution of ^{1}H chemical shifts (having the two 180° pulses does not affect J evolution).

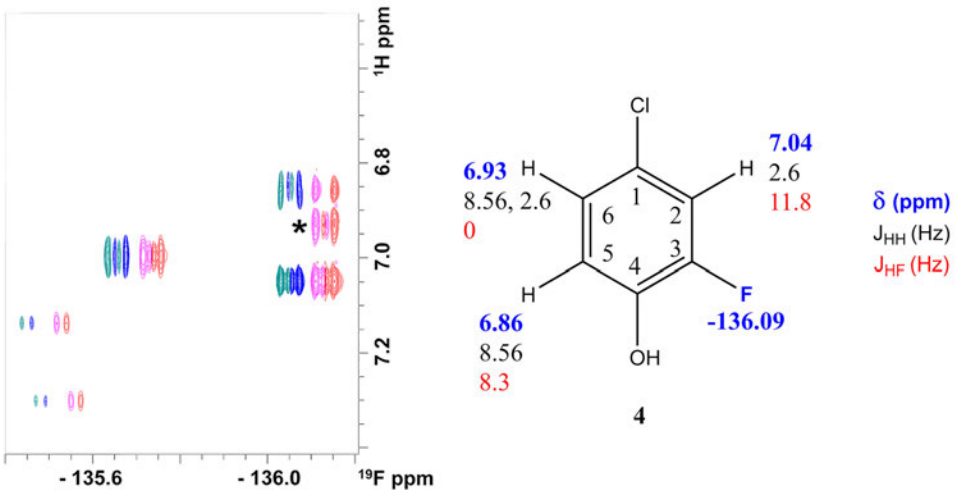


Figure 6.9: Overlay of 2D ^{1}H - ^{19}F HETCOR spectrum (blue/green) and 2D ^{1}H - ^{19}F TOCSY-HETCOR (red/pink) highlighting the need/usefulness of the TOCSY element. Note red/pink cross-peaks have been shifted to the right to facilitate the inspection of the spectra.

As can be seen in Figure 6.9, a by-product of this reaction (**4**) appears to have a proton within the same spin system as the fluorine, which does not directly couple to it. This is likely a result of the chlorine in position 1, manipulating the electron density within the ring. It is the TOCSY element within the 2D ^{1}H - ^{19}F TOCSY-HETCOR sequence that allows its information to be obtained.

6.4.2.3 *1D and 2D ^{1}H - ^{19}F Cross Polarisation (CP) experiment*

The cross peaks in the 2D ^{1}H - ^{19}F HMBC and TOCSY-HETCOR spectra are pure phase proton multiples in F1, while HETCOR shows mixed phase multiplets due to J_{HH} evolution during the Δ_1 + Δ_2 interval. However, the typical t_1 acquisition times in these experiments are too short to resolve the ^{1}H - ^{1}H couplings and these are best measured in the directly detected dimension. This is straightforward in ^{1}H - detected ^{19}F HETCOR, but not with the ^{1}H, ^{19}F HMBC, as this experiment yields mixed phase proton multiplets. In its basic form, the F2 multiplets in ^{1}H detected, ^{1}H - ^{19}F HETCOR have an antiphase proton-fluorine doublet component with additional in phase proton-proton

splitting. However, the antiphase nature of these multiplets can complicate the process of coupling constant determination.

For simple mixtures of fluorinated compounds, the proton-proton couplings can be measured efficiently from in phase multiplets by a selective-reverse ^{19}F, ^{1}H INEPT that is followed by a proton-TOCSY step as illustrated recently.[146] Here the proton-fluorine coupling is refocused (which requires the resonance frequency of a fluorine coupled proton to be known and to be amenable to selective inversion) prior to the ^{1}H - ^{1}H TOCSY transfer. Consequently, fluorine decoupling can be used to produce pure in phase multiplets split only by the proton-proton couplings.

For complex mixtures, this experiment is impractical; a different approach is required, which does not require selective excitation of fluorine resonances and selective inversion of a particular proton. An NMR building block that meets these criteria is the ^{19}F - ^{1}H cross-polarization (CP) as its in-phase nature of the magnetisation transfer (condition to a suitably selected mixing time) allows immediate application of ^{19}F decoupling. An experiment bearing these attributes, a 3D CP ^{19}F - ^{1}H heteronuclear TOCSY, has been designed for resonance assignment of fluorinated compounds.[147] A 1D version of this experiment is presented in Figure 6.10. The flow of magnetisation is described in equation 6.3.

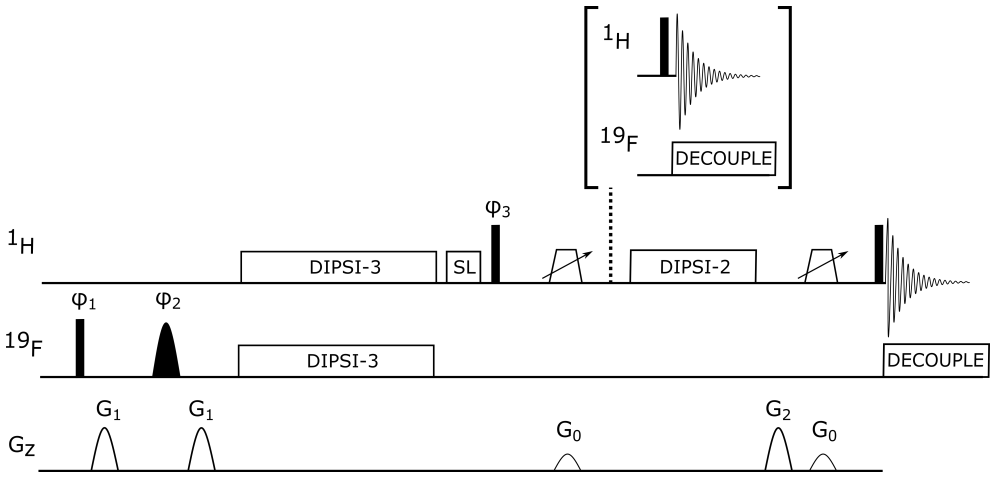


Figure 6.10: Pulse sequence of 1D ^{19}F - ^{1}H CP experiment, the insert represented with a dotted line shows an alternative arrangement, where additional proton-proton spin lock is not applied after the cross-polarisation step. Thin filled rectangles represent high power 90° pulses, the filled sine shaped pulse with represents a selective 180° pulse, the pulse labelled "SL" represents a spin-lock pulse and the trapezoid pulse with an inclined arrow represents a 180° CHIRP pulse. The spin lock pulse, DIPSI-3, DIPSI-2 spin locks and the 90° pulse following the spin lock (φ_3) were applied at $\gamma B_1/2\pi$ = 6.25 KHz (PW_{90} = 40 μs) on both ^{1}H and ^{19}F . The following phases were used: φ_1 = y, -y; φ_2 = 2x, 2y; φ_3 = 4y, 4(-y) and the receiver phase = x, 2(-x), x, -x, 2x, -x. Gradient strengths are as follows: G_0 = 5%; G_1 = 10% and G_2 = 7%

The short version of this 1D CP sequence (indicated by the dashed line) without an additional ^{1}H spin lock was used to determine the excitation bandwidth of several spin lock schemes, two of which, DIPSI3 and FLOPSY-16 cycle,[148,149] are shown below (Figure 6.11). This was performed by shifting the ^{19}F carrier frequency in increments of 1 ppm, whilst retaining the same carrier frequency for the selective ^{19}F pulse. As can be seen in Figure 6.11 the FLOPSY-16 marginally outperforms the DIPSI3 cycle. Considering offset where at least some spin-lock was effective, FLOPSY-16 excites ± 6031 Hz, which on a 400 MHz spectrometer equates to a total width of 32 ppm, whereas the DIPSI3 only excites ± 5277 Hz, or a 28 ppm window.

$$^{19}\mathrm{F} \xrightarrow[\text{DIPSI-3}]{^{n}J_{HF}} \xrightarrow[\text{DIPSI2}]{^{n}J_{HH}} {}^{1}\mathrm{H}(t_2) \tag{6.3}$$

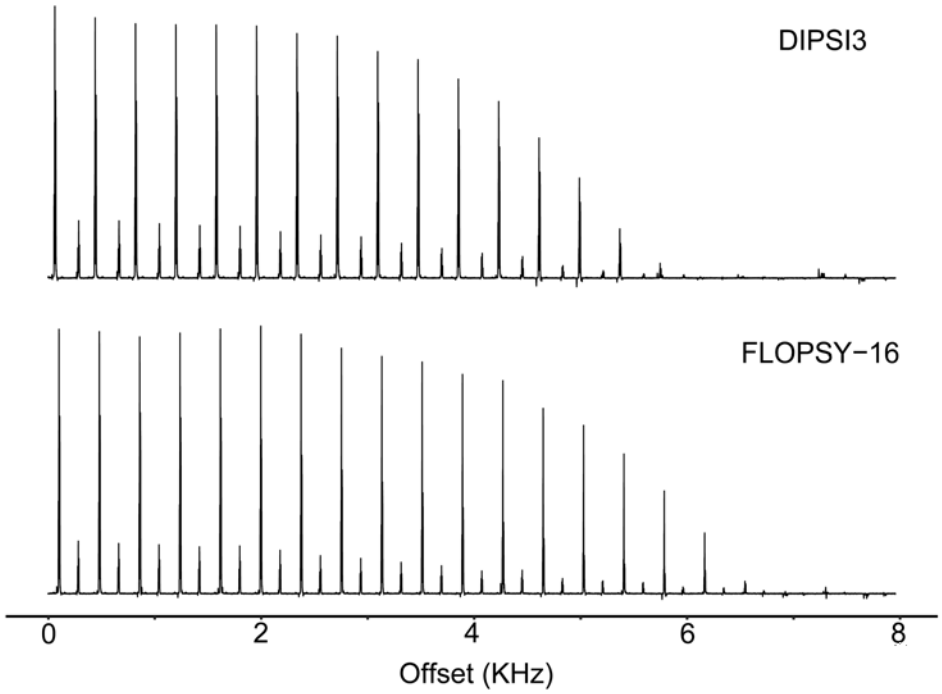


Figure 6.11: Excitation profiles for DIPSI3 and FLOPSY-16 spin locks, produced using the 1D ^{19}F - ^{1}H CP experiment.

This 1D sequence is useful if the chemical shift of the fluorine is already known and their number is limited. However, when analysing complex, fluorine containing mixtures, multiple species are required to be analysed at the same time. For this we have developed a 2D version (Figure 6.12) of the 3D sequence proposed in[147]. Unlike in the 3D sequence, we did not find it necessary to label the ^{1}H chemical shifts after the initial transfer from ^{19}F. The flow of magnetisation in the 2D CP experiment is shown in equation 6.4.

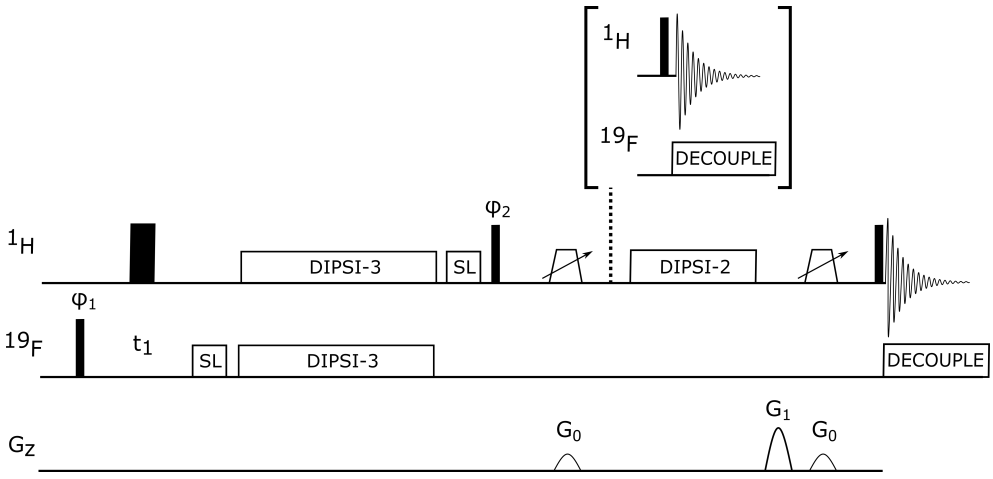


Figure 6.12: Pulse sequence of 2D ^{19}F - ^{1}H CP experiment, the insert represented with a dotted line shows an alternative arrangement, where additional proton-proton spin lock is not applied after the cross-polarisation step. Thin and thick filled rectangles represent high power 90° and 180° pulses respectively, the filled sine shaped pulse with represents a selective 180° pulse, the pulse labelled "SL" represents a spin-lock pulse and the trapezoid pulse with an inclined arrow represents a 180° CHIRP pulse. The spin lock pulse, DIPSI-3, DIPSI-2 spin locks and the 90° pulse following the spin lock (φ_3) were applied at $\gamma B_1/2\pi$ = 6.25 KHz (PW_{90} = 40 μs) on both ^{1}H and ^{19}F . The following phases were used: φ_1 = y, -y; φ_2 = 2y, 2(-y) and the receiver phase = x, 2(-x), x. Gradient strengths are as follows: G_0 = 5% and G_1 = 7%

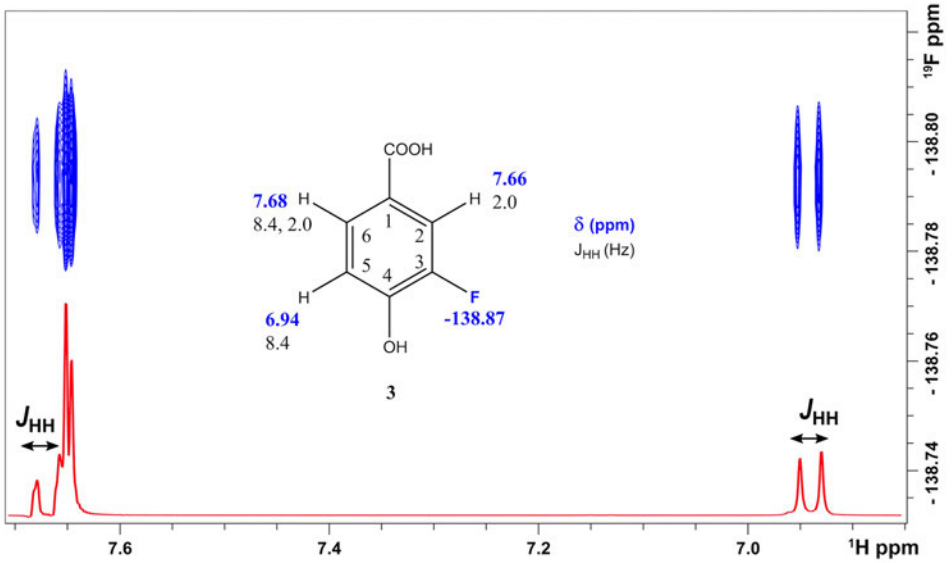


Figure 6.13: Spectrum of 2D ^{19}F - ^{1}H CP experiment, highlighting cross peaks from **3**. The 1D trace shown in red was taken from the 2D spectrum and highlights the in-phase nature of the J_{HH} .

Optionally, the signal in these (1D and 2D) CP ^{19}F - ^{1}H experiments can be acquired after the initial CP step and z-filter.[150] As the proton-proton spin-lock acts already during the CP step, magnetisation of some of the protons not coupled to ^{19}F is already available. Alternatively, a ^{1}H DIPSI-2 step can be added to extend the transfer to further parts of the spin system and spread the magnetisation more evenly. Application of two z-filters and the ^{19}F decoupling ensures that pure in phase multiplets split only by proton-proton couplings are acquired. Figure 6.13, shows the cross peaks of protons in compound **3**, allowing for the determination of J_{HH}, although there is significant overlap of protons 2 and 6.

$$^{19}\mathrm{F} \xrightarrow{t_1(^{19}\mathrm{F})} \xrightarrow[\text{DIPSI-3}]{^{n}J_{HF}} \xrightarrow[\text{DIPSI-2}]{^{n}J_{HH}} {}^{1}\mathrm{H}(t_2) \tag{6.4}$$

6.4.2.4 *One-bond and long-range 2D ^{19}F - ^{13}C HMBC*

The structure determination of sparsely protonated fluorinated molecules, such as heavily substituted aromatic rings, based only on ^{1}H and ^{19}F chemical shifts is problematic. Nevertheless, thanks to far reaching ^{19}F, ^{13}C couplings ($^{n}J_{FC}$, where n = 1-5), many ^{13}C chemical shifts can be obtained via 2D ^{19}F - ^{13}C correlated experiments such as HMBC or HSQC that make their structure determination possible. As the $^{1}J_{FC}$ couplings are large ($\approx$ 150-250 Hz), while the $^{n>1}$J$_{FC}$ typically range from 0 to 35 Hz,[151] the one-bond and long-range correlation experiments are best optimised separately. A single "long-range" experiment can fortuitously also yield one-bond correlations if it happens that multiple rotations of ^{19}F magnetisation during the evolution interval fall outside of the even multiple of $1/^{1}J_{FC}$. This can only be achieved on purpose if the values of the coupling constants are known and their spread is narrow.

The 2D ^{19}F - ^{13}C HMBC experiments optimised for one-bond and long-range correlation are presented in Figure 6.14.

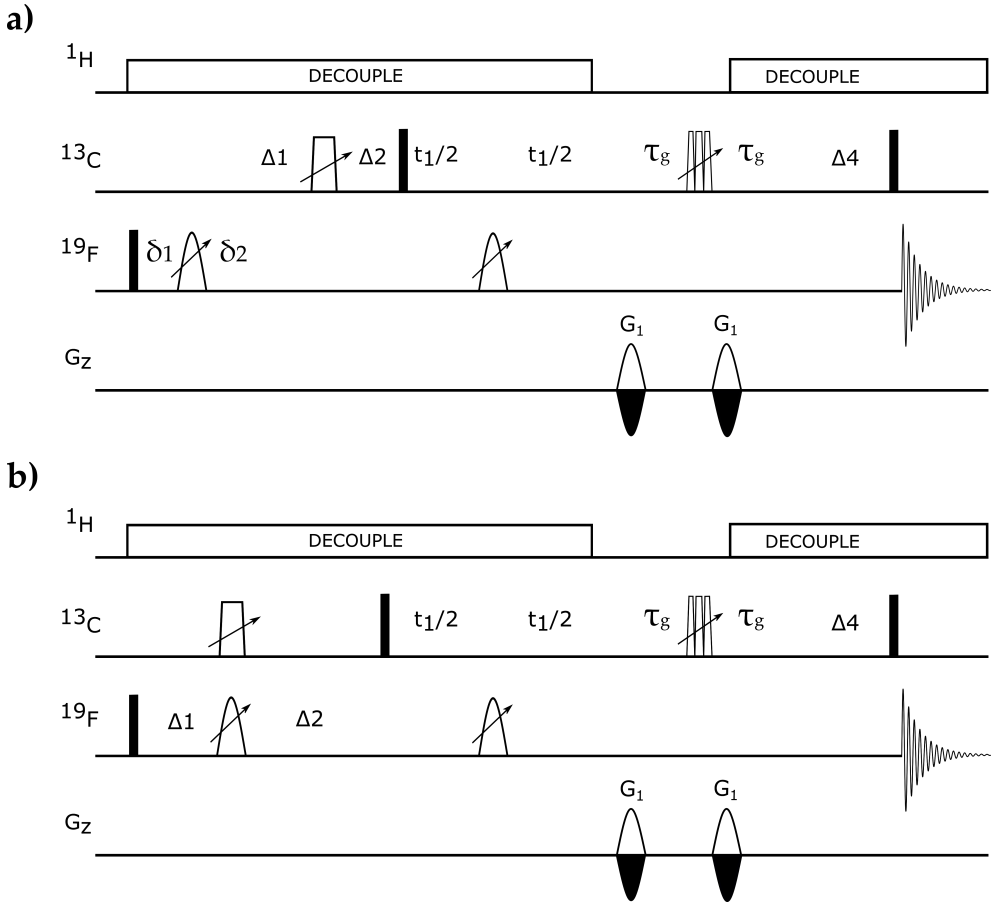


Figure 6.14: Pulse sequences of 2D ^{19}F - ^{13}C HMBC optimised for **a)** $^{1}J_{FC}$ and **b)** $^{n}J_{FC}$ correlations. The thin and thick filled rectangles represent high power 90° and 180° pulses. The ^{19}F adiabatic CHIRP pulses (p12, 1 ms) and inversion (p14, 500 μs) and refocusing (p24, 2 ms) ^{13}C CHIRP pulses are indicated by an inclined arrow. Unless stated otherwise, pulses were applied from the x axis.

In **a)** the following delays were used: δ_1 = 20 μs, δ_2 = (2/π) p1 + 20 μs, Δ_1 = (Δ_3/2) – d6 – (p14/2), Δ_2 = (Δ_3/2) + d6 – (p14/2) – p3, Δ_3 = 2τ_g + p24 + Δ_4, Δ_4 = p12 + t$_1$(0). In **b)** the following delays were used: Δ_1 < 1/4 J_{FC} (optimised for 20 Hz in this work), Δ_2 = Δ_1 – p14 + (2/π) p1 + 2τ_g + Δ_4 + p24 - p3, Δ_4 = p12 + t$_1$(0), and Δ_3 = 2τ_g + p24 + Δ_4, where d6 = 1/4 $^{1}J_{FC}$, p1 is the 90° ^{19}F pulse, p3 is the 90° ^{13}C pulse, τ_g is the PFG + the recovery delay (1.2 ms) and t$_1$(0) is the initial t$_1$ evolution delay (6 μs).

The following phases were used: φ_1 = 2x, 2(-x); φ_2 = x, -x; φ_3 = 4x, 4(-x) and the receiver phase = 2(x,-x), 2(-x, x). Echo-anti-echo protocol was used for sign discrimination in F1 by changing the polarity of the PFGs. [152]

When these sequences were initially developed, ^{1}H decoupling was only implemented during acquisition, the result of this was still the suppression of J_{HF} couplings in F2, but without ^{1}H decoupling throughout the length of the experiment, the J_{HC} couplings could still evolve resulting in cross peaks being split in F1. This reduces signal intensity (due to more splitting), but allows quick discovery of those carbons which have directly attached protons as seen in Figure 6.15b. By employing the ^{1}H decoupling throughout the entirety of the experiment, the cross peaks are substantially simplified to singlets in F1 and antiphase doublets in F2.

Similarly to the 2D ^{1}H - ^{19}F HMBC, magnetization is not refocused in F2 due to the spread of the J_{FC} sizes. The J_{FC} couplings are present as pure phase antiphase doublets in F2 providing valuable structural information. An example spectrum is shown in Figure 6.15a.

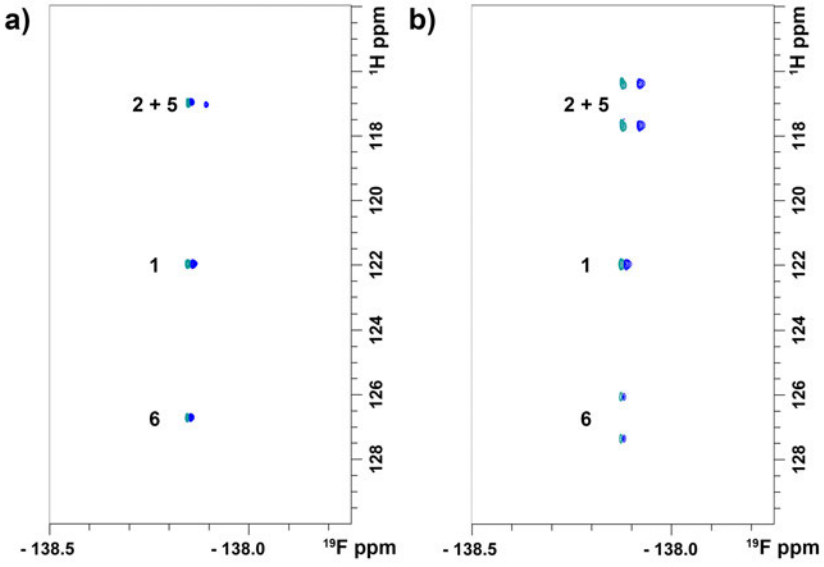
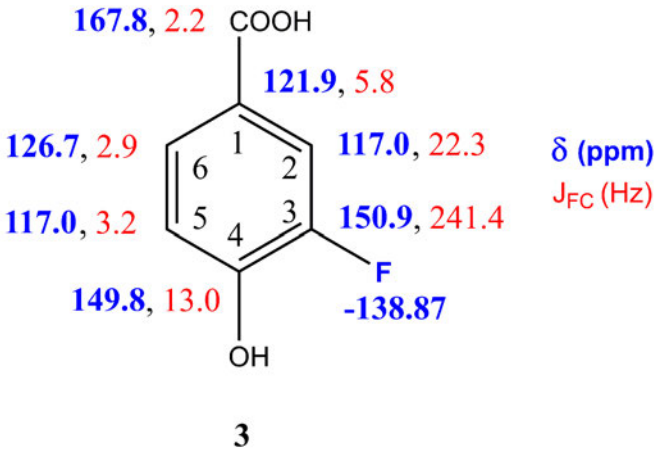


Figure 6.15: Spectrum of the ^{19}F - ^{13}C HMBC experiment optimised to $^{n}J_{FC}$ correlations. The experiment provides ^{19}F - ^{13}C coupling constants in antiphase. Numbers to the left of cross peaks indicate the assigned carbon (based on assignments shown in Figure 6.16).

$$\begin{aligned} &\text{a) } {}^{19}\text{F} \xrightarrow{{}^{1}J_{FC}} \xrightarrow{t_1({}^{13}\text{C})} {}^{19}\text{F}(t_2) \\ &\text{b) } {}^{19}\text{F} \xrightarrow{{}^{n}J_{FC}} \xrightarrow{t_1({}^{13}\text{C})} {}^{19}\text{F}(t_2) \end{aligned} \tag{6.5}$$



3

Figure 6.16: Structure of the starting material (**3**) showing the ^{19}F / ^{13}C **chemical shifts (δ, ppm)** and the ^{19}F - ^{13}C coupling constants.

The 2D ^{19}F - ^{13}C HMBC experiments, allow for the acquisition of the chemical shifts of those carbons J coupled to a fluorine. The coupling constants are incredibly useful for determining the position of the carbon relative to the fluorine.

6.4.2.5 *Reduced dimensionality (3,2)D ^{1}H - ^{13}C - ^{19}F correlation experiment*

One-bond ^{1}H - ^{13}C correlation is a cornerstone of the structural determination of compounds by NMR. For mixtures of fluorinated compounds, it is essential that this information is extended and combined with the ^{1}H - ^{19}F and ^{13}C - ^{19}F correlations provided by the experiments described above. This is best achieved by correlating all three types of nuclei in a HCF experiment. Inspired by the 3D HNCO experiment, one of the protein back bone assignment pulse sequences, a 3D triple-resonance ^{1}H , ^{13}C , ^{19}F experiment has been proposed.[153,154] This out-and-back 3D ^{1}H - detected method is performed with ^{13}C and ^{19}F decoupling and contains $^{n}J_{FC}$ defocusing and refocusing intervals. So far, it has been applied to fluorinated polymers with a narrow dispersion of ^{19}F chemical shifts.

An implementation of the HCF experiment using a direct polarisation transfer pathway ^{1}H $\rightarrow$ ^{13}C $\rightarrow$ ^{19}F in the form of a reduced dimensionality (3,2)D HCF experiment (Figure 6.17) is presented here.[155] In this experiment, the ^{1}H chemical shift is recorded first, while suppressing the evolution of ^{1}H - ^{1}H and ^{1}H - ^{19}F coupling constants by a BIRD^{rX} pulse[156,157] and a 180° ^{19}F pulse in the middle of the t_1 period, respectively. The magnetisation is then transferred via $^{1}J_{CH}$ couplings to carbons in an INEPT step, where is it refocused before starting ^{1}H decoupling. During this interval, the ^{19}F

- ^{13}C antiphase magnetisation is allowed to develop. At the same time, the central 180° ^{13}C and 180° ^{19}F pulses are moved simultaneously with the t_1 incrementation, modulating the ^{1}H chemical shifts by ^{13}C chemical shift and splitting the cross peaks into F1 doublets. The signal is then transferred to ^{19}F, where it is detected in the presence of ^{1}H decoupling as antiphase ^{19}F - ^{13}C doublets. The F1 doublets carry the ^{13}C chemical shift information. Interleaved acquisition of two spectra differing by a 90° phase shift of the φ_2 phase allows spectra to be simplified. As described previously in[158] and[159], the resulting in phase or antiphase doublets can be edited to remove one line of the doublet at a time. Spectra are acquired using an echo-antiecho protocol and accompanying pulsed field gradients.

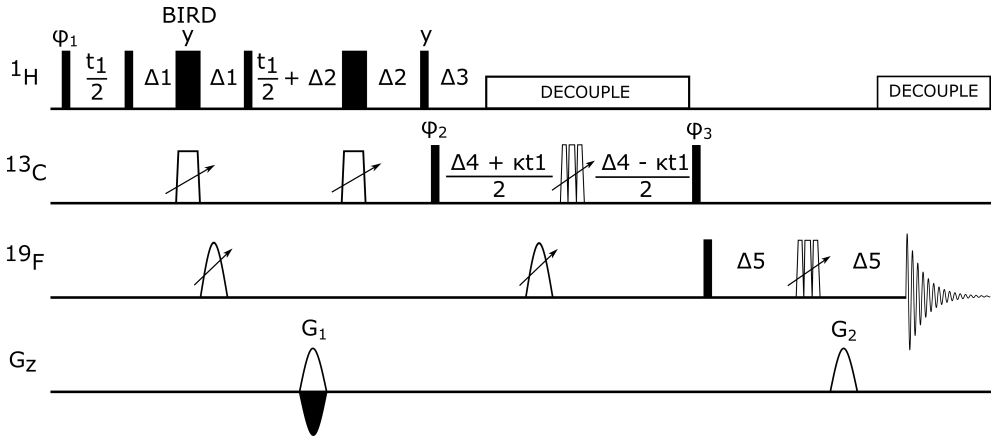


Figure 6.17: Reduced dimensionality (3, 2)D HCF correlation experiment. Narrow and wide filled rectangles represent 90° and 180° pulses, respectively. Trapezoids with inclined arrows represent 500 μs ^{13}C CHIRP pulses and 2 ms composite 180° ^{13}C CHIRP pulses. Sine shaped pulses with inclined arrows represent 180° 1 ms adiabatic pulses.

The following delays were used: Δ_1 = (1/2 $^{1}J_{CH}$) - (p14/2); Δ_2 = (1/4 $^{1}J_{CH}$) - (p14/2); Δ_3 = (1/2 $^{1}J_{CH}$) for CH and (1/4 * $^{1}J_{CH}$) for all multiplicities; Δ_4 = (1/2 J_{FC}); Δ_5 = 1.2 ms (1ms PFG and 200 μs recovery delay); κ is the scaling factor for $\Omega_{^{13}C}$ frequencies. Unless specified otherwise, pulses were applied from the x axis; φ_1 = y, -y; φ_2 = 4x, 4(-x); φ_3 = 2x, 2(-x) or 2y, 2(-y) for cosine and sine modulated signal respectively; the receiver phase = x, 2(-x), x, -x, 2x, -x. The real and imaginary points were acquired by changing the polarity of the G_1 gradient. G_1 = 80% and G_2 = 85.04%, where 100% represents 53 Gauss/cm.

The resulting 2D spectrum contains information about the chemical shifts of all three nuclei and also $^{n}J_{FC}$ coupling constants, as shown for the starting material (**3**) in Figure 6.18. Inclusion of only one $^{n}J_{FC}$ evolution interval

provides more flexibility in the range of couplings mediating the transfer, while the modulation of the signals by the ^{13}C chemical shifts has an equivalent effect on the obtainable signals to noise as an extension of the experiment to three dimensions. High resolution can be achieved using the variable-time t_1 period of this 2D experiment without crosspeak broadening. The ^{13}C shift modulation can be scaled relative to the t_1 evolution (κ factor), keeping the F1 spectral width small without limitations on the length of the constant-time $^{n}J_{FC}$ evolution interval, Δ_4.

$$^{1}H \xrightarrow{t_1(^{1}H)} \xrightarrow{^{1}J_{HC}} \xrightarrow{\kappa t_1(^{13}C)} \xrightarrow{^{n}J_{FC}} {}^{19}F(t_2) \tag{6.6}$$

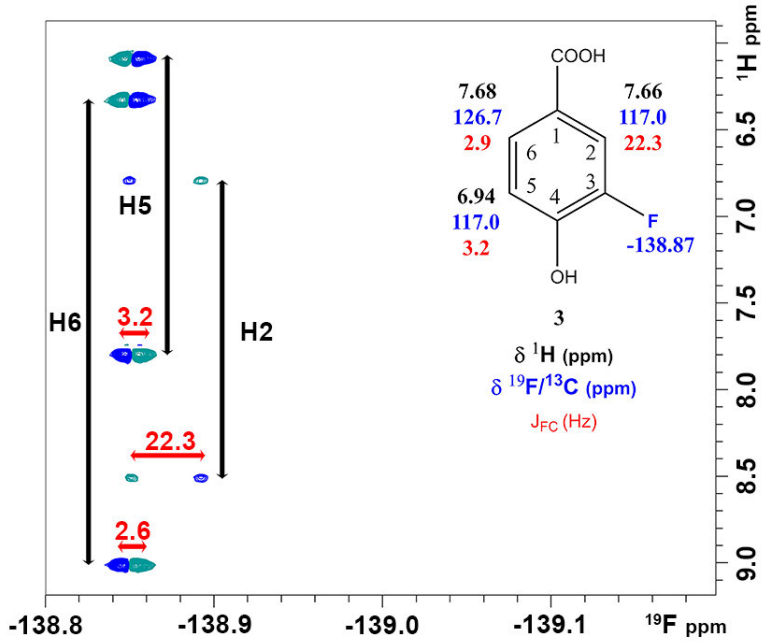


Figure 6.18: HCF spectrum showing signals from the starting material (**3**). Black arrows indicate the splitting of proton cross peaks, the middle of which is the true proton chemical shift. Red arrows represent J_{FC} for the carbons directly bonded to the respective protons.

To obtain the proton chemical shifts from Figure 6.18, the mid-point of the F1 doublets must be taken. In order to procure the ^{13}C chemical shifts directly bonded to these protons, equation 6.7 must be used. Where O2p is the carrier frequency of ^{13}C , Δ is the splitting of the doublet in Hz, K is the scaling factor set by the user and SF ^{13}C is the spectrometer frequency for carbon. An example of this is given in equation 6.7b for carbon 2 of compound **3**.

$$\begin{aligned} &\text{a) } \delta\ ^{13}\text{C}_{\text{ ppm}} = \text{O2p} + \frac{\Delta(\text{Hz})}{2 \times K \times \text{SF}\ ^{13}\text{C}} \\ &\text{b) } 117\ \text{ppm} = 100 + \frac{859}{2 \times 0.2 \times 125.77} \end{aligned} \tag{6.7}$$

In this chapter, a set of NMR experiments for the analysis of mixtures of fluorinated compounds was presented. In the following a protocol explaining how these experiments come together to derive the structures of individual molecules is presented and applied to the mixture produced by the chloramination of compound **3**.

6.4.3 *Protocol for the analysis of monofluorinated mixtures*

In order to effectively characterise these species using some or all of the above described experiments, a general protocol was devised and laid out in Figure 6.19.

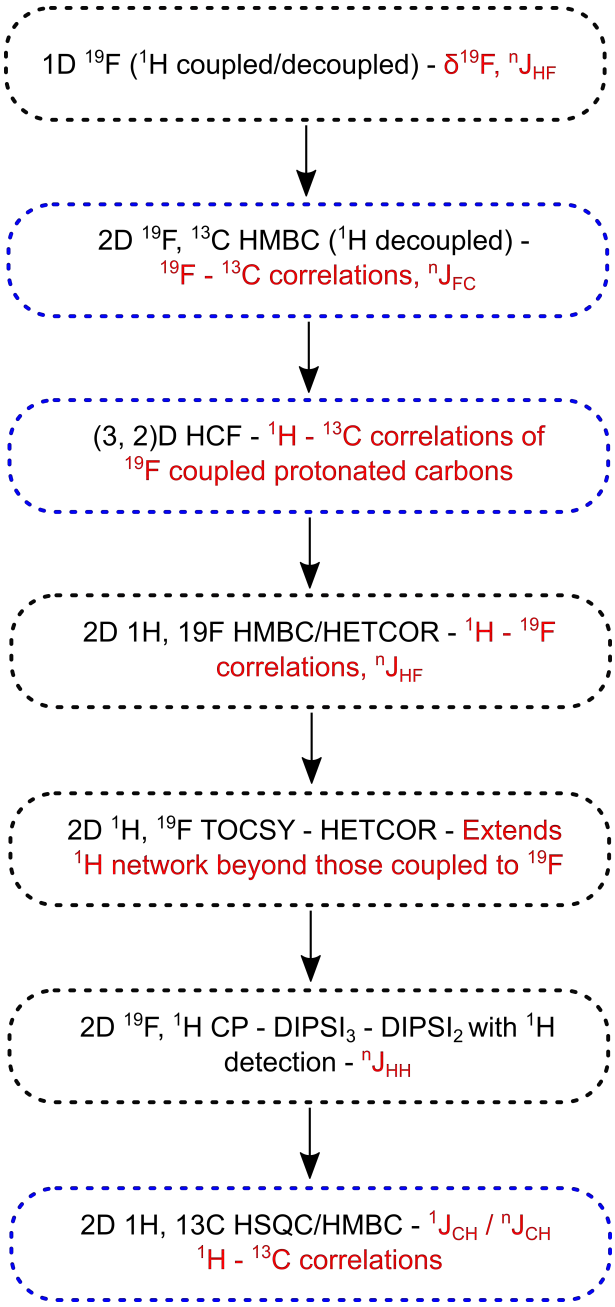


Figure 6.19: Flow chart for acquiring and working with spectra of complex fluorinated mixtures. Information obtained from the given experiment is given in red. Blue boxes denote experiments involving ^{13}C.

This flow chart summarises the information that is obtained from individual experiments, together with the recommended order to run them. If the compound concentration is low, it might be difficult to obtain spectral information from those experiments that utilise ^{13}C; these experiments are highlighted in blue.

6.4.4 *Structure elucidation of the chloramination by-products of compound 3*

The first step in analysing these complex fluorinated mixtures, is to acquire 1D ^{19}F spectra, both ^{1}H-decoupled and ^{1}H-coupled, as shown in Figure 6.20. Decoupling increases the sensitivity of the experiment, providing the best possibility of obtaining the chemical shifts of those by-products produced in lower concentrations. However, the coupled experiment allows for the determination of the sizes of ^{19}F - ^{1}H coupling constants; these are often crucial in the elucidation of small molecules.

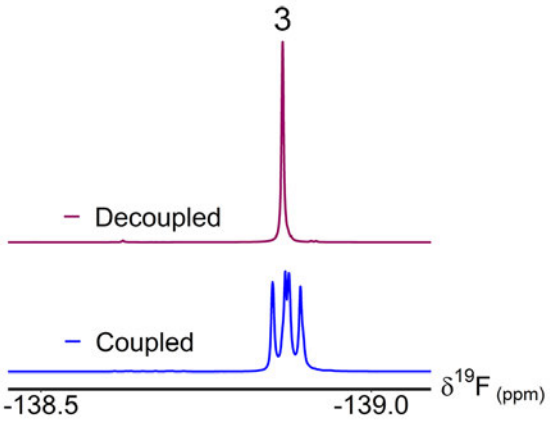


Figure 6.20: 1D ^{19}F experiment with (purple) and without (blue) ^{1}H decoupling.

Inspection of the ^{1}H decoupled 1D ^{19}F experiment shows that there is a significant concentration of the starting material still present. Its ^{19}F signal appears at -138.87 ppm. The ^{19}F - ^{13}C HMBC indicates this ^{19}F shares a J_{FC} coupling (241.4 Hz) with an *ipso* aromatic ^{13}C, with a ^{13}C δ of 150.854 ppm. The ^{19}F also shares J_{FC} couplings with δ ^{13}C = 117.0 (o, 22.3 Hz), 116.98 (p), 126.71 (m, 2.9 Hz) and 149.79 (o, 13.2 Hz), 167.83 (m, 2.2 Hz) ppm. The ^{19}F - ^{1}H HCF experiment indicates the F shares J_{HF} couplings with protons at 6.94 (8.9 Hz), 7.65 (11.5 Hz), 7.68 (0.8 Hz) ppm. The J_{HH} indicated in brackets were determined accurately using 1D traces from the CP - DIPSI3 - DIPSI2 experiment. Examining the size of the J_{FC} and J_{HF} couplings indicates the following ^{13}C/^{1}H correlation at 117.0/7.65 ppm, confirmed by the ^{1}H - ^{13}C HSQC. The ^{1}H - ^{13}C HMBC confirmed the order of the carbons in the aromatic ring. Combining the data, indicates that the intense signal at -138.87 ppm is indeed molecule **3**, the starting material (Figure 6.21).

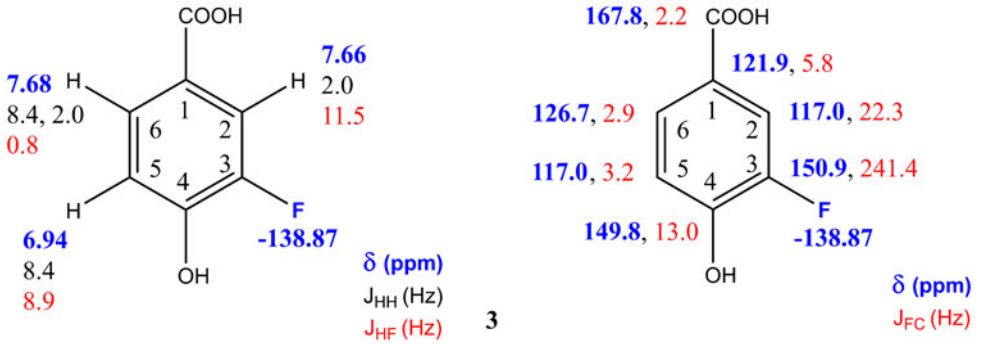


Figure 6.21: Structure of the starting material (**3**) highlighting the spectral information gained through the use of the designed ^{19}F correlated experiments.

The only by-product of a comparable concentration to the starting material is compound **5**; its structure differs from the starting material only by the addition of a chlorine. Whilst investigating its structure, a strange phenomenon was noticed: An additional signal (split by J_{HF} couplings) at the level of the ^{13}C satellites appeared in the 1D ^{19}F spectra, close to the signal of a one-bond ^{13}C satellite. It did not originate from interactions with other spin $\frac{1}{2}$ nuclei, as its paired signal was missing (Figure 6.22). Its appearance was attributed to solvent effects. This sample was dissolved in CD_3OH, however, there is some residual non-deuterated solvent (CH_3OH). This causes a solvent induced isotope shift, which in the literature is quoted as being approximately 0.2 ppm.[160,161]

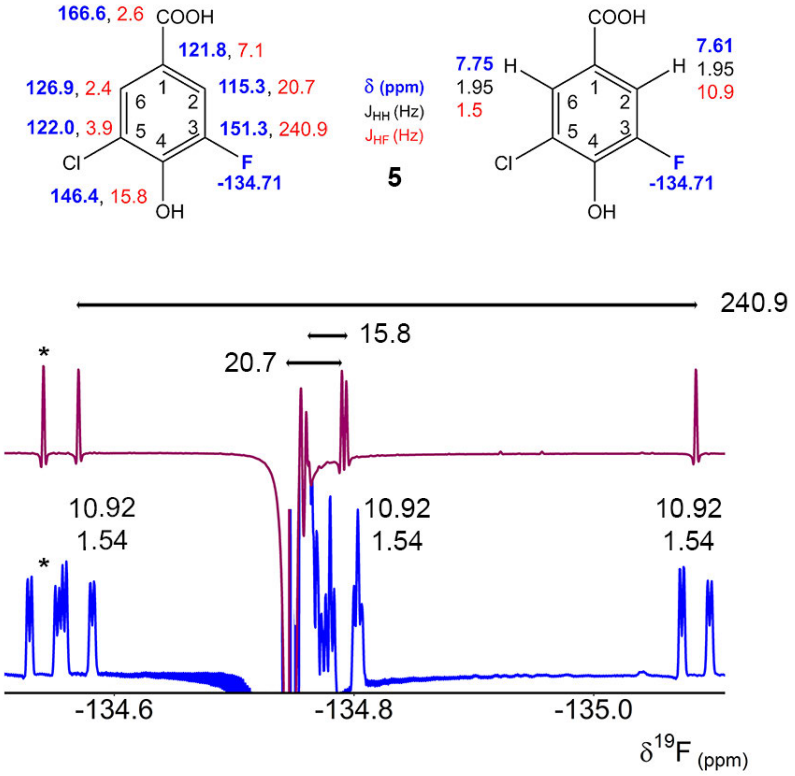


Figure 6.22: 1D ^{19}F spectra without (blue) and with (purple) ^{1}H decoupling, highlighting the ^{13}C satellites of compound **5**. * denotes the signal resulting from the solvent-induced isotope shift of the main ^{12}C isotopomer.

At first glance, the spectrum does not appear to be all that complex, however, a closer inspection shows that there are many species of a lower concentration with ^{19}F signals resonating between -125 and -140 ppm (Figure 6.23).

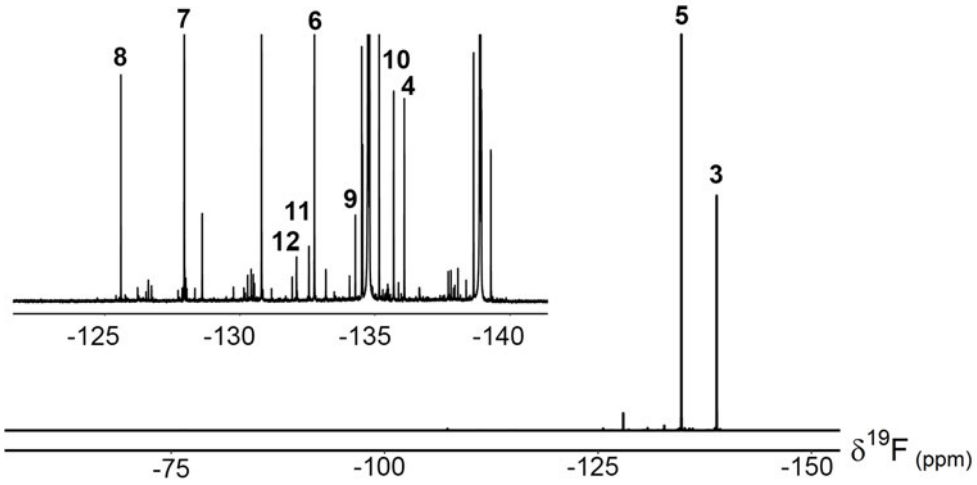


Figure 6.23: ^{1}H decoupled 1D ^{19}F experiment. Spectrum in the top left of the figure is a zoomed region for displaying of lower concentration compounds. Bold numbers indicate compound assignments.

Outside of this region, signals were identified that show large ^{19}F - ^{15}N coupling constants (Figure 6.24). Their presence was confirmed by ^{15}N decoupling. The coupling constants of this magnitude(around 50 Hz) have been reported before in systems like those shown in Figure 6.24, where the large coupling is a result of through space interactions.[162,163] Such large couplings suggest the presence of larger, constrained systems, produced perhaps through oligomerisation.

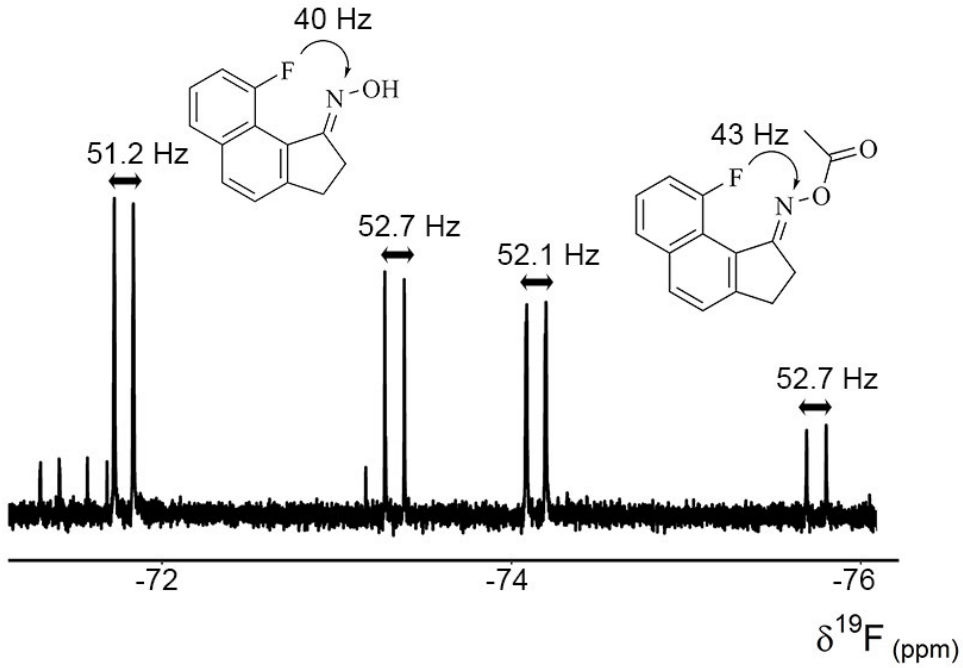


Figure 6.24: 1D ^{19}F NMR spectrum with ^{1}H decoupling. Doublets shown are species which have large ^{19}F - ^{15}N coupling constants. Compounds shown are examples taken from the following references[162] and[163].

The mechanisms for the production of by-products through chloramination are not fully understood, however, there are some studies that have monitored their production.[164] For example, a reaction pathway between aldehydes and monochloramine is presented in Figure 6.25, adapted from[164].

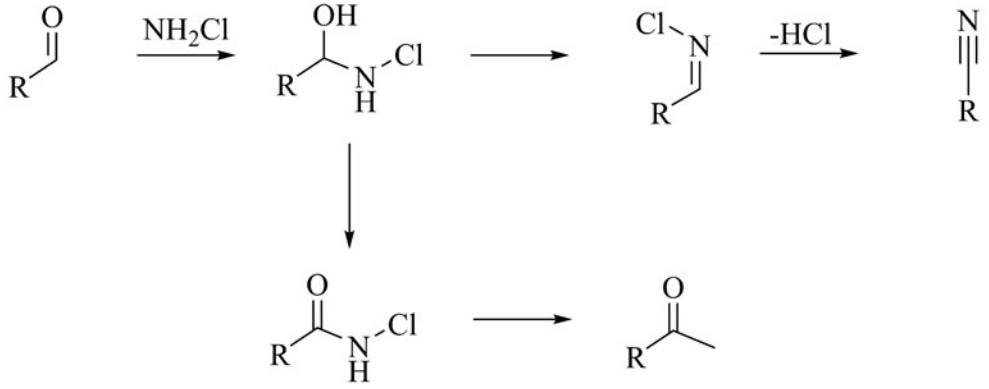


Figure 6.25: Reaction pathway of monochloramine (NH_2Cl) with an aldehyde, forming organic chloramines, nitriles and amides. Adapted from[164].

Figure 6.26 shows the by-products produced from the chloramination of **3**, which contain a single aromatic ring similar to the starting material. Compound **5**, is the most intense, as seen in Figure 6.23, this represents halogenation

(substitution reaction) of the starting material, *ortho* to the hydroxyl group, in position 5. Halogenation via decarboxylation is also possible producing compound **4**, this mechanism combined with the first resulted in a dichlorinated product in compound **6**.

The mechanism by which compound **7** is produced from the starting material is not immediately obvious, but there is no trace of this compound in any of the blanks performed throughout this experiment. This product is then also chlorinated, *ortho* to the hydroxyl group in position 6, to produce compound **8**.

Compound **9**, is produced in an analogous way to compound **5**, but brominated instead of chlorinated. Bromine is naturally present in sodium hypochlorite,[54] and is likely to be present as hypobromous acid, which, like hypochlorous acid can go on and react with DOM to produce brominated by-products.[165–167]

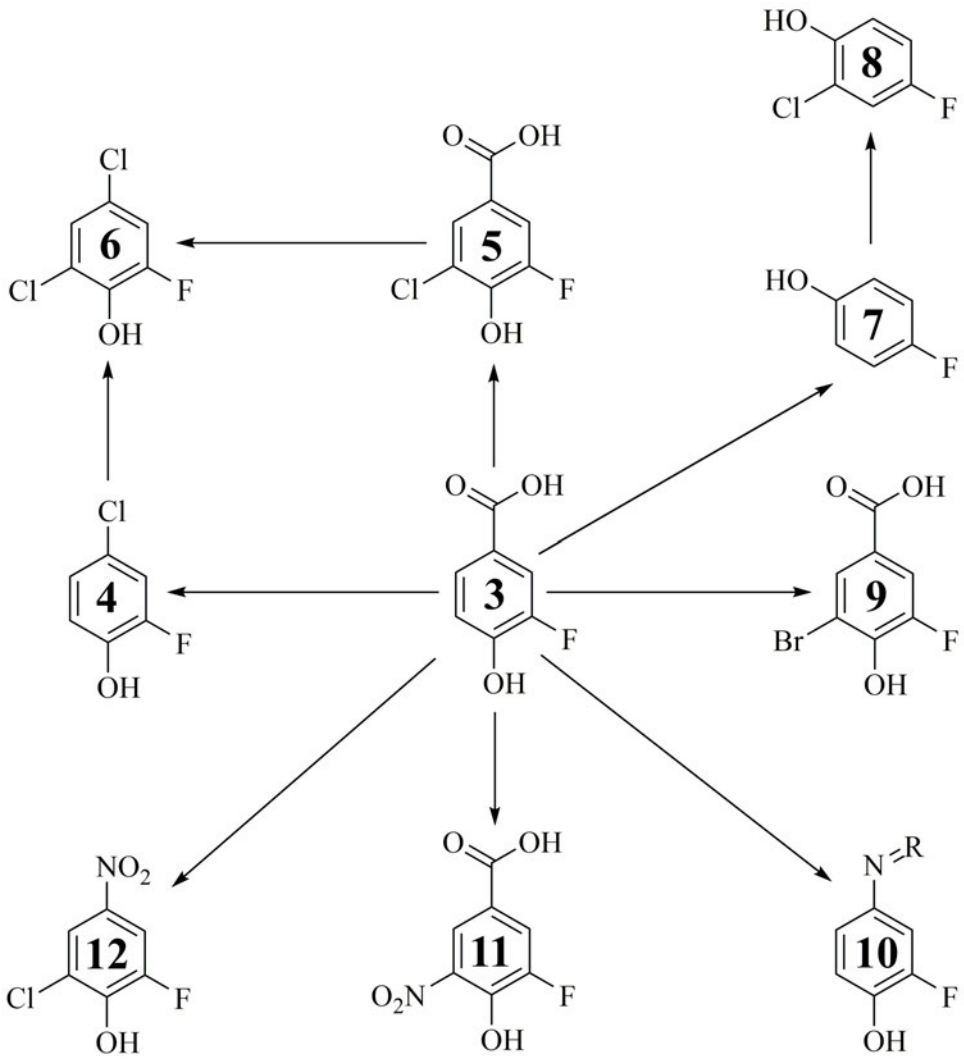


Figure 6.26: Starting material (**3**) and all identified by-products.

Compounds **10**, **11** and **12** all show evidence of coupling to another NMR active nucleus, beyond ^{1}H and ^{13}C. This was deemed to be nitrogen and confirmed through the use of $^{15}N/^{1}H$ - decoupled ^{19}F spectra as shown in Figure 6.27 for compound **10** showing a ^{19}F - ^{15}N coupling of 1.5 Hz. Spectral data provided for structures such as 4-hydroxy-3-nitrobenzoic acid,[168] suggest that it is a nitro group in compounds **11** and **12**. The nature of the nitrogen substituent in compound **10** could not be identified.

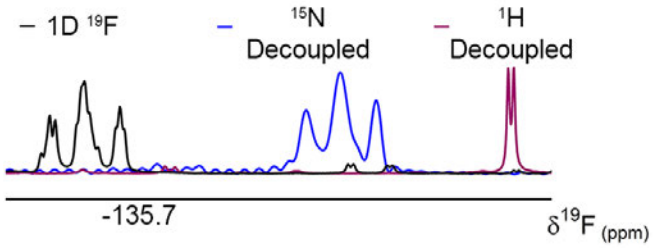


Figure 6.27: 1D ^{19}F spectra without decoupling (black) with ^{15}N decoupling (blue) and with ^{1}H decoupling, highlighting the presence of ^{15}N on compound **10**.

NMR parameters for the by-products **4** - **12** can be found in Appendix Chapter D, along with descriptions of how the structures of these molecules were determined.

6.5 CONCLUSIONS

Chloramination of DOM showed a significant incorporation of nitrogen into SRFA, a DOM standard. When model compounds, small molecules such as **1** and **2**, were chloraminated, they also produced a complex mixture of DBPs. This demonstrated that this reaction can produce hundreds of species that can overwhelm analytical methods trying to characterise them. A strategy was therefore designed to allow the structural determination of major species by using a fluorinated precursor, 3-Fluoro-4-hydroxybenzoic acid.

This lead us to optimise existing and develop new NMR experiments for the interrogation or probing of complex mixtures of small molecules containing a single ^{19}F nucleus. These experiments allowed for the correlation of ^{19}F, ^{1}H and ^{13}C chemical shifts. They also provide values of the J_{FC} and J_{HF} coupling constants. This allowed for the assignment of a dozen compounds within the complex mixture, without the need for physical separation processes.

The fluorinated mixture used in this work was the product of the chloramination reaction across 5 days in the dark, producing one major chlorinated species,

a dozen or so other compounds approximately $\frac{1}{30}$ the intensity of the major by-product, and hundreds of compounds below 1%.

Those DBPs produced in the largest concentration were mostly chlorinated without sign of nitrogen addition. Some lower concentration compounds were found to have incorporated ^{15}N. The ^{1}H - ^{15}N correlation experiments indicated that many more are present than indicated in the ^{19}F spectra.

CONCLUSIONS AND FUTURE WORK

Throughout this project two very powerful, high-resolution analytical techniques, NMR and FT-ICR-MS, were used to characterise dissolved organic matter (DOM), a highly complex, heterogeneous mixture that can interfere with most stages of water treatment.

In this work, two small-scale water treatment works (pilot plants) were investigated. One utilising a novel suspended ion-exchange (SIX) and CeraMac (ceramic) membrane filtration, the other comparing ion-exchange (IEX) and granulated activated carbon (GAC) filtration as post coagulation treatments. Using FT-ICR-MS, NMR and ATR-FTIR it was found that the SIX treatment was non-selective in its removal of DOM, whereas the ceramic membrane filtration was selectively targeting more aromatic and phenolic species. This work suggests that if SIX is capable of removing DOM indiscriminately, i.e. can remove those species that coagulation/filtration struggles with, it may be better suited as a post coagulation treatment, eliminating DOM concentrations even further, prior to the disinfection process.

Whilst investigating the second pilot plant, it was found that of the two treatment methods investigated, the IEX filtration outperformed the GAC over a 6 month period. The IEX treatment resulted in a greater reduction of total organic carbon (TOC), reduced the number of molecular formulae identified and a greater reduction in aromatic ^{1}H NMR intensity. Aromatic and phenolic species have been shown to be precursors of trihalomethanes; their reduction should result in a lower production of regulated compounds. The methodology used in the studies of these pilot plants provides tools that go beyond standard analytical techniques currently in place to assess water quality and the molecular level information represents a valuable addition to water providers.

Advanced oxidation processes (AOPs) are said to be the treatment methods of the 21st century, capable of efficiently degrading organic material. Photocatalysis is an AOP that is gaining a lot of traction, with many studies investigating its use in water treatment. In this work, a study comparing the most popular photocatalyst, TiO_2, to a TiO_2 based catalyst modified with bismuth titanate

(BTO - TiO_2) was performed. NMR and FT-ICR-MS were used to characterise the composition of DOM over a 3 hour period of photocatalytic degradation with different irradiation wavelengths. Suwannee river fulvic acid (SRFA), an internationally recognised DOM standard, was used as a model DOM in this study. It was found, by FT-ICR-MS, that under the shortest wavelength investigated, 370 nm , the two catalysts performed similarly with comparable levels of degradation. However, the efficiencies started to diverge with the longer wavelengths investigated, i.e. 410 nm and a white LED (430 nm +). The modified catalyst, due to the presence of a heterojunction is capable of utilising lower energy wavelengths to excite electrons; this was reflected in the degradation of DOM, with more species being lost completely at these wavelengths. Due to increased efficiency at longer wavelengths, a modified catalyst can reduce the demand for higher energy UV lamps. ^{1}H NMR was also used and confirmed the conclusions drawn from the MS data, this is reassuring as NMR is a quantitative technique that reports higher concentration compounds. Principle components analysis (PCA) of ^{1}H NMR spectra showed that the major products of photocatalysis were long chain fatty acids. This study represents the most detailed molecular level characterisation of the photocatalytic degradation of DOM to date. The techniques for evaluating molecular formulae, derived from FT-ICR-MS spectra, developed here will help to assess the efficiency of new catalysts and monitor their performance in the degradation of complex mixtures.

A major part of potable water production is disinfection, without which the water would be unfit for consumption. Chlorination, the most common method of disinfection, results in the production of chlorinated disinfection by-products (DBPs), some of which are regulated. In order to minimise the production of these regulated by-products, alternative disinfection methods are becoming increasingly more popular. In this work chloramination is investigated at the molecular level. It is known to produce less regulated DBPs, however we showed via ^{15}N NMR, that there is significant incorporation of nitrogen into DOM. An initially surprising observation was that hundreds of molecules were also produced when individual, small aromatic acids were chloraminated. This initial observation, lead us to the study of a model system, a small molecule that incorporated fluorine as a convenient NMR tag. A suite of ^{19}F based NMR experiments were developed to take advantage of this sensitive nucleus. These experiments provided ^{19}F, ^{1}H and ^{13}C chemical shifts, as well as J_{FC} and J_{HF} coupling constants. Using this information, structures of several by-products were determined in a complex mixture without the need for their physical separation. This methodology will be used to study the reaction

between disinfectants with model compounds, their kinetics and efficiency, assisting the development of water disinfection procedures. It has the potential to contribute to the production of higher quality potable water.

BIBLIOGRAPHY

(1) J. D'Andrilli, C. M. Foreman, A. G. Marshall and D. M. McKnight, *Organic Geochemistry*, 2013, **65**, 19 –28.

(2) S. A. Huber, A. Balz, M. Abert and W. Pronk, *Water Research*, 2011, **45**, 879 –885.

(3) G. Riise, P. V. Hees, U. Lundström and L. T. Strand, *Geoderma*, 2000, **94**, 237 –247.

(4) M. E. Essington, *Soil and Water Chemistry - An Integrative Approach*, CRC Press, 2003.

(5) A. Piccolo, *Adv. Environ. Res*, 1999, **3**, 508–521.

(6) A. Piccolo, *Soil Science*, 2001, **166**, 810–832.

(7) J. A. Hawkes, P. J. R. Sjöberg, J. Bergquist and L. J. Tranvik, *Faraday Discuss.*, 2019, **218**, 52–71.

(8) S. Wang and C. N. Mulligan, *Chemosphere*, 2009, **74**, 274–279.

(9) A. A. Helal, D. Imam and S. Khalifa, *Radiochemistry*, 2006, **48**, 419–425.

(10) R. Hahn, C. Hein, J. M. Sander and R. Kautenburger, *Applied Geochemistry*, 2017, **78**, 241 –249.

(11) Drinking Water Quality in Scotland 2018 Public Water Supply - Annual Report, *DWQR*, 2018.

(12) P. Jarvis, B. Jefferson and S. Parsons, *Water Supply*, 2004, **4**, 79–87.

(13) E. Sharp, P. Jarvis, S. Parsons and B. Jefferson, *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 2006, **286**, 104 –111.

(14) J. E. V. Benschoten and J. K. Edzwald, *Water Research*, 1990, **24**, 1527 –1535.

(15) G. Korshin, C. W. Chow, R. Fabris and M. Drikas, *Water Research*, 2009, **43**, 1541 –1548.

(16) T. Li, Z. Zhu, D. Wang, C. Yao and H. Tang, *Powder Technology*, 2006, **168**, 104 –110.

(17) E. L. Sharp, S. A. Parsons and B. Jefferson, *Environmental Pollution*, 2006, **140**, 436 –443.

(18) G. C. Budd, A. F. Hess, H. Shorney-Darby, J. J. Neemann, C. M. Spencer, J. D. Bellamy and P. H. Hargette, *American Water Works Association*, 2004, **96**, 102–113.

(19) K. Bell-Ajy, M. Abbaszadegan, E. Ibrahim, D. Verges and M. LeChevallier, *American Water Works Association*, 2000, **92**, 44–58.

(20) J. Yu, D. Wang, M. Yan, C. Ye, M. Yang and X. Ge, *Envronmental Monitoring and Assessment*, 2007, **131**, 377 –386.

(21) P. Jarvis, E. Sharp, M. Pidou, R. Molinder, S. A. Parsons and B. Jefferson, *Water Research*, 2012, **46**, 4179 –4187.

(22) D. Lakshmanan, D. Clifford and G. Samanta, *American Water Works Association*, 2008, **100**, 76–88.

(23) P. Jarvis, J. Banks, R. Molinder, T. Stephenson, S. A. Parsons and B. Jefferson, *Water Supply*, 2008, **8**, 709–716.

(24) E. Sharp, J. Banks, J. Billica, K. Gertig, R. Henderson, S. Parsons, D. Wilson and B. Jefferson, *Water Supply*, 2005, **5**, 49–56.

(25) A. Morfesis, A. M. Jacobson, R. Frollini, M. Helgeson, J. Billica and K. R. Gertig, *Industrial & Engineering Chemistry Research*, 2009, **48**, 2305–2308.

(26) F. S. Cannon, V. L. Snoeyink, R. G. Lee, G. Dagois and J. R. DeWolfe, *American Water Works Association*, 1993, **85**, 76–89.

(27) A. M. El-Rehaili and W. J. Weber, *Water Research*, 1987, **21**, 573 –582.

(28) J. J. McCreary and V. L. Snoeyink, *Water Research*, 1980, **14**, 151 –160.

(29) F. Li, A. Yuasa, K. Ebie, Y. Azuma, T. Hagishita and Y. Matsui, *Water Research*, 2002, **36**, 4592 –4604.

(30) B. Bolto, D. Dixon, R. Eldridge and S. King, *Water Research*, 2002, **36**, 5066 –5073.

(31) J. Morran, M. Drikas, D. Cook and D. Bursill, *Water Supply*, 2004, **4**, 129–137.

(32) P. C. Singer and K. Bilyk, *Water Research*, 2002, **36**, 4009 –4022.

(33) P. Jarvis, M. Mergen, J. Banks, B. McIntosh, S. A. Parsons and B. Jefferson, *Environmental Science & Technology*, 2008, **42**, 1276–1282.

(34) G. Galjaard, E. Koreman, D. Metcalfe, G. Moore and P. Ericsson, *Water Practice and Technology*, 2018, **13**, 524–541.

(35) L. Verdickt, W. Closset, V. D'Haeseleer and J. Cromphout, *Water Supply*, 2012, **12**, 878–887.

(36) Ángeles de la Rubia, M. Rodríguez, V. M. León and Daniel, *Water Research*, 2008, **42**, 714 –722.

(37) L. Sun, E. Perdue and J. McCarthy, *Water Research*, 1995, **29**, 1471 –1477.

(38) J. Koprivnjak, E. Perdue and P. Pfromm, *Water Research*, 2006, **40**, 3385 –3392.

(39) T. Vetter, E. Perdue, E. Ingall, J.-F. Koprivnjak and P. Pfromm, *Separation and Purification Technology*, 2007, **56**, 383 –387.

(40) S. Nakagawara, T. Goto, M. Nara, Y. Ozawa, K. Hotta and Y. Arata, *Analytical Sciences*, 1998, **14**, 691–698.

(41) W. C. Barrette, D. M. Hannum, W. D. Wheeler and J. K. Hurst, *Biochemistry*, 1989, **28**, 9172–9178.

(42) J. C. Morris, *The Journal of Physical Chemistry*, 1966, **70**, 3798–3805.

(43) G. A. Mcfeters and A. K. Camper, in, ed. A. I. Laskin, Academic Press, 1983, vol. 29, pp. 177 –193.

(44) J. M. Albrich, C. A. McCarthy and J. K. Hurst, *Proceedings of the National Academy of Sciences*, 1981, **78**, 210–214.

(45) S. M. McKenna and K. J. A. Davies, *Biochemical Journal*, 1988, **254**, 685–692.

(46) S. D. Richardson, M. J. Plewa, E. D. Wagner, R. Schoeny and D. M. DeMarini, *Mutation Research/Reviews in Mutation Research*, 2007, **636**, 178 –242.

(47) J Rook, *Water Treatment Examination*, 1974, **23**, 234 –243.

(48) T. Bellar, J. Lichtenberg and R. Kroner, *American Water Works Association*, 1974, **66**, 703–706.

(49) A. D. Shah and W. A. Mitch, *Environmental Science & Technology*, 2012, **46**, 119–131.

(50) M. Deborde and U. von Gunten, *Water Research*, 2008, **42**, 13 –51.

(51) G. Bitton, Microbiology of Drinking Water Production and Distribution, *John Wiley & Sons*, 2014, 298.

(52) R. E. Raudales, J. L. Parke, C. L. Guy and P. R. Fisher, *Agricultural Water Management*, 2014, **143**, 9 –28.

(53) Y. E. Lin, J. E. Stout and V. L. Yu, *Infection Control & Hospital Epidemiology*, 2011, **32**, 166–173.

(54) *World Health Organisation*, Guidelines for drinking-water quality: fourth edition incorporating the first addendum, 2017, 336.

(55) H. Sakai, S. Tokuhara, M. Murakami, K. Kosaka, K. Oguma and S. Takizawa, *Water Research*, 2016, **88**, 661 –670.

(56) W.-H. Chen and T. M. Young, *Environmental Science & Technology*, 2008, **42**, 1072–1077.

(57) R. Homlok, E. Takács and L. Wojnárovits, *Chemosphere*, 2013, **91**, 383 –389.

(58) S. Goel, R. M. Hozalski and E. J. Bouwer, *Journal - AWWA*, 1995, **87**, 90–105.

(59) J Rositano, G Newcombe, B Nicholson and P Sztajnbok, *Water Research*, 2001, **35**, 23 –32.

(60) N. Nakada, H. Shinohara, A. Murata, K. Kiri, S. Managaki, N. Sato and H. Takada, *Water Research*, 2007, **41**, 4373 –4382.

(61) H. J. H. Fenton, *Journal of the Chemical Society, Transactions*, 1894, **65**, 899–910.

(62) Y. Nie, C. Hu, J. Qu and X. Hu, *Journal of Hazardous Materials*, 2008, **154**, 146 –152.

(63) S.-Y. Pang, J. Jiang and J. Ma, *Environmental Science & Technology*, 2011, **45**, 307–312.

(64) A. Khataee, P. Gholami and M. Sheydaei, *Journal of the Taiwan Institute of Chemical Engineers*, 2016, **58**, 366 –373.

(65) A. I. Zárate-Guzmán, L. V. González-Gutiérrez, L. A. Godínez, A. Medel-Reyes, F. Carrasco-Marín and L. A. Romero-Cano, *Chemosphere*, 2019, **224**, 698 –706.

(66) S. Navalon, M. Alvaro and H. Garcia, *Applied Catalysis B: Environmental*, 2010, **99**, 1 –26.

(67) T. Mackuľak, M. Mosný, R. Grabic, O. Golovko, O. Koba and L. Birošová, *Environmental Toxicology and Pharmacology*, 2015, **39**, 483 –488.

(68) G. T. Chi, J. Churchley and K. D. Huddersman, *International Journal of Chemical Engineering*, 2013, **2013**, ed. F. Fatone, 760915.

(69) C. A. Murray and S. A. Parsons, *Chemosphere*, 2004, **54**, 1017 –1023.

(70) J. J. Molnar, J. R. Agbaba, B. D. Dalmacija, S. D. Rončević, S. P. Maletić and M. M. Kragulj, *Journal of Advanced Oxidation Technologies*, 2011, **14**, 54 –62.

(71) Sanly, M. Lim, K. Chiang, R. Amal, R. Fabris, C. Chow and M. Drikas, *Separation Science and Technology*, 2007, **42**, 1391–1404.

(72) M. Trapido and J. Kallas, *Environmental Technology*, 2000, **21**, 799–808.

(73) M. D. Gurol and R. Vatistas, *Water Research*, 1987, **21**, 895 –900.

(74) N. Takahashi, *Ozone: Science & Engineering*, 1990, **12**, 1–17.

(75) M. Trapido, A. Hirvonen, Y. Veressinina, J. Hentunen and R. Munter, *Ozone: Science & Engineering*, 1997, **19**, 75–96.

(76) A. Hirvonen, T. Tuhkanen and P. Kalliokoski, *Water Science and Technology*, 1996, **33**, 67–73.

(77) M. Ortega-Liébana, E. Sánchez-López, J. Hidalgo-Carrillo, A. Marinas, J. Marinas and F. Urbano, *Applied Catalysis B: Environmental*, 2012, **127**, 316 –322.

(78) D. Hermosilla, M. Cortijo and C. P. Huang, *Science of the Total Environment*, 2009, **407**, 3473–3481.

(79) W. Carmen, A Comprehensive Dictionary of Inorganic Chemistry, *Abhishek*, 2009.

(80) M. R. Hoffmann, S. T. Martin, W. Choi and D. W. Bahnemann, *Chemical Reviews*, 1995, **95**, 69–96.

(81) G. Al-Sayyed, J.-C. D'Oliveira and P. Pichat, *Journal of Photochemistry and Photobiology A: Chemistry*, 1991, **58**, 99 –114.

(82) S. S. Alias, Z. Harun, F. H. Azhar, S. A. Ibrahim and B. Johar, *Journal of Cleaner Production*, 2020, **251**, 119448.

(83) M. Zulfiqar, S. Chowdhury, S. Sufian and A. A. Omar, *Journal of Cleaner Production*, 2018, **203**, 848 –859.

(84) R. Khunphonoi and N. Grisdanurak, *Chemical Engineering Journal*, 2016, **296**, 420 –427.

(85) P. Hore, *Nuclear Magnetic Resonance*, Oxford University Press, 2nd edn., 2015.

(86) D. Jeannerat, Rapid Multidimensional NMR: High Resolution by Spectral Aliasing, *eMagRes*, 2011.

(87) G. L. Bretthorst, *Concepts in Magnetic Resonance Part A*, 2008, **32A**, 417–435.

(88) A. J. Simpson, *Magnetic Resonance in Chemistry*, 2002, **40**, S72–S82.

(89) A. J. Simpson, B. Lefebvre, A. Moser, A. Williams, N. Larin, M. Kvasha, W. L. Kingery and B. Kelleher, *Magnetic Resonance in Chemistry*, 2004, **42**, 14–22.

(90) J. Buddrus, P. Burba, J. Lambert and H. Herzog, *Analytical Chemistry*, 1989, **61**, 628–631.

(91) E. B. Ledford, D. L. Rempel and M. L. Gross, *Analytical Chemistry*, 1984, **56**, 2744–2748.

(92) E. N. Nikolaev, R. M. A. Heeren, A. M. Popov, A. V. Pozdneev and K. S. Chingin, *Rapid Communications in Mass Spectrometry*, 2007, **21**, 3527–3546.

(93) J. A. Loo, H. R. Udseth and R. D. Smith, *Analytical Biochemistry*, 1989, **179**, 404 –412.

(94) P. Kebarle and U. H. Verkerk, *Mass Spectrometry Reviews*, 2009, **28**, 898–917.

(95) M. Glückmann, A. Pfenninger, R. Krüger, M. Thierolf, M. Karasa, V. Horneffer, F. Hillenkamp and K. Strupat, *International Journal of Mass Spectrometry*, 2001, **210-211**, 121 –132.

(96) M. Karas and R. Krüger, *Chemical Reviews*, 2003, **103**, 427–440.

(97) J. A. Barry, G. Robichaud and D. C. Muddiman, *Journal of the American Society for Mass Spectrometry*, 2013, **24**, 1137–1145.

(98) W. Haas, B. K. Faherty, S. A. Gerber, J. E. Elias, S. A. Beausoleil, C. E. Bakalarski, X. Li, J. Villén and S. P. Gygi, *Molecular & Cellular Proteomics*, 2006, **5**, 1326–1337.

(99) D. F. Smith, A. Kharchenko, M. Konijnenburg, I. Klinkert, L. Paša-Tolić and R. M. A. Heeren, *Journal of The American Society for Mass Spectrometry*, 2012, **23**, 1865–1872.

(100) B. M. Kujawinski EB, *Analytical Chemistry*, 2006, **78**, 4363–4373.

(101) D. P. A. Kilgour, C. L. Mackay, P. R. R. Langridge-Smith and P. B. O'Connor, *Analytical Chemistry*, 2012, **84**, 7431–7435.

(102) D. V. Krevelen, *Fuel*, 1950, **29**, 269–284.

(103) A. Rivas-Ubach, Y. Liu, T. S. Bianchi, N. Tolić, C. Jansson and L. Paša-Tolić, *Analytical Chemistry*, 2018, **90**, 6152–6160.

(104) A. Kellerman, T. Dittmar, D. Kothwala and L. Tranvik, *Nature Communications*, 2014, **5**, 3804.

(105) I. Douterelo, S. Husband, V. Loza and J. Boxall, *Applied and Environmental Microbiology*, 2016, **82**, ed. J. L. Schottel, 4155–4168.

(106) W. Kew, J. W. Blackburn, D. J. Clarke and D. Uhrín, *Rapid Communications in Mass Spectrometry*, 2017, **31**, 658–662.

(107) D. M. Osborne, D. C. Podgorski, D. A. Bronk, Q. Roberts, R. E. Sipler, D. Austin, J. S. Bays and W. T. Cooper, *Rapid Communications in Mass Spectrometry*, 2013, **27**, 851–858.

(108) T. Dittmar, B. Koch, N. Hertkorn and G. Kattner, *Limnology and Oceanography: Methods*, 2008, **6**, 230–235.

(109) A. Lex, N. Gehlenborg, H. Strobelt, R. Vuillemot and H. Pfister, *IEEE Transactions on Visualization and Computer Graphics*, 2014, **20**, 1983–1992.

(110) Z. Yuan, C. He, Q. Shi, C. Xu, Z. Li, C. Wang, H. Zhao and J. Ni, *Environmental Science & Technology*, 2017, **51**, 8110–8118.

(111) E. Lavonen, D. Kothawala, L. Tranvik, M. Gonsior, P. Schmitt-Kopplin and S. Köhler, *Water Research*, 2015, **85**, 286 –294.

(112) D. Metcalfe, C. Rockey, B. Jefferson, S. Judd and P. Jarvis, *Water Research*, 2015, **87**, 20 –28.

(113) T. Bond, E. Goslan, S. Parsons and B. Jefferson, *Environmental Technology*, 2011, **32**, 1–25.

(114) B. Lam and A. J. Simpson, *Analyst*, 2008, **133**, 263–269.

(115) R. R. Artz, S. J. Chapman, A. J. Robertson], J. M. Potts, F. Laggoun-Défarge, S. Gogo, L. Comont, J.-R. Disnar and A.-J. Francez, *Soil Biology and Biochemistry*, 2008, **40**, 515 –527.

(116) C. Cocozza, V. D'Orazio, T. Miano and W. Shotyk, *Organic Geochemistry*, 2003, **34**, 49 –60.

(117) D. Gondar, R. Lopez, S. Fiol, J. Antelo and F. Arce, *Geoderma*, 2005, **126**, 367 –374.

(118) P. Zaccheo, G. Cabassi, G. Ricca and L. Crippa, *Organic Geochemistry*, 2002, **33**, 327 –345.

(119) M. Deborde and U. [von Gunten], *Water Research*, 2008, **42**, 13 –51.

(120) S. D. Boyce and J. F. Hornig, *Environmental Science & Technology*, 1983, **17**, 202–211.

(121) N. Tolić, Y. Liu, A. Liyu, Y. Shen, M. M. Tfaily, E. B. Kujawinski, K. Longnecker, L.-J. Kuo, E. W. Robinson, L. Paša-Tolić and N. J. Hess, *Analytical Chemistry*, 2017, **89**, 12659–12665.

(122) A. Matilainen, N. Lindqvist, S. Korhonen and T. Tuhkanen, *Environment International*, 2002, **28**, 457 –465.

(123) M. Valdivia-Garcia, 2017, Trihalomethanes: From Precursors to Management Strategies, PhD Thesis, Newcastle University, Newcastle.

(124) G. Odling, Z. Y. Pong, G. Gilfillan, C. R. Pulham and N. Robertson, *Environ. Sci.: Water Res. Technol.*, 2018, **4**, 2170–2178.

(125) G. Odling, E. Chatzisymeon and N. Robertson, *Catal. Sci. Technol.*, 2018, **8**, 829–839.

(126) E. M. Perdue, Functions of Natural Organic Matter in Changing Environment, ed. J. Xu, J. Wu and Y. He, Springer Netherlands, Dordrecht, 2013, pp. 85–88.

(127) X. Huang, M. Leal and Q. Li, *Water research*, 2008, **42**, 1142—1150.

(128) *Water Research*, 2017, **125**, 201 –208.

(129) S. Liu, M. Lim, R. Fabris, C. Chow, M. Drikas and R. Amal, *Organic Geochemistry*, 2010, **41**, 124 –129.

(130) *Water Research*, 2006, **40**, 323 –330.

(131) L. A. T. Espinoza], E. [ter Haseborg], M. Weber and F. H. Frimmel, *Applied Catalysis B: Environmental*, 2009, **87**, 56 –62.

(132) G. Odling and N. Robertson, *ChemSusChem*, 2015, **8**, 1838–1840.

(133) X. Huang, M. Leal and Q. Li, *Water research*, 2008, **42**, 1142—1150.

(134) A. Assabane, Y. Ait Ichou, H. Tahiri, C. Guillard and J.-M. Herrmann, *Applied Catalysis B: Environmental*, 2000, **24**, 71 –87.

(135) B. Kraeutler and A. J. Bard, *Journal of the American Chemical Society*, 1978, **100**, 5985–5992.

(136) E. Bracco, M. Butler and Carnelli, *Environmental Science and Pollution Research*, 2020, **27**, 28469–28479.

(137) M. G. Muellner, E. D. Wagner, K. McCalla, S. D. Richardson, Y.-T. Woo and M. J. Plewa, *Environmental Science & Technology*, 2007, **41**, 645–651.

(138) H. Barjat, G. Morris, S. Smart, A. Swanson and S. Williams, *Journal of Magnetic Resonance, Series B*, 1995, **108**, 170 –172.

(139) N. Bell, A. Michalchuk, J. Blackburn, M. Graham and D. Uhrín, *Angewandte Chemie*, 2015, **54**, 8382 –8385.

(140) H. Zhang and M. Yang, *Science of The Total Environment*, 2018, **627**, 118 –124.

(141) X. Wang, J. Wang, Y. Zhang, Q. Shi, H. Zhang, Y. Zhang and M. Yang, *Science of The Total Environment*, 2016, **554-555**, 83 –88.

(142) A. S. Ginwalla and M. A. Mikita, *Environmental Science & Technology*, 1992, **26**, 1148–1150.

(143) L. E. Kay, M. Ikura, R. Tschudin and A. Bax, *Journal of Magnetic Resonance*, 1990, **89**, 496–514.

(144) J. Battiste and R. Newmark, *Progress in Nuclear Magnetic Resonance Spectroscopy*, 2006, **48**, 1–23.

(145) J. Bohlen and G. Bodenhausen, *Journal of Magnetic Resonance, Series A*, 1993, **102**, 293 –301.

(146) L. Castañar, P. Moutzouri, T. M. Barbosa, C. F. Tormena, R. Rittner, A. R. Phillips, S. R. Coombes, M. Nilsson and G. A. Morris, *Analytical Chemistry*, 2018, **90**, 5445–5450.

(147) H. Hu, P. Kulanthaivel and K. Krishnamurthy, *The Journal of Organic Chemistry*, 2007, **72**, 6259–6262.

(148) A. Shaka, C. Lee and A Pines, *Journal of Magnetic Resonance (1969)*, 1988, **77**, 274 –293.

(149) M Kadkhodaie, O Rivas, M Tan, A Mohebbi and A. Shaka, *Journal of Magnetic Resonance (1969)*, 1991, **91**, 437 –443.

(150) M. J. Thrippleton and J. Keeler, *Angewandte Chemie International Edition*, 2003, **42**, 3938–3941.

(151) Guide to Fluorine NMR for Organic Chemists, chapter 3, *John Wiley & Sons*, 2016, 55–132.

(152) D. O. Cicero, G. Barbato and R. Bazzo, *Journal of Magnetic Resonance*, 2001, **148**, 209 –213.

(153) L. Li and P. L. Rinaldi, *Macromolecules*, 1996, **29**, 4808–4810.

(154) L. Li, B. Zhang, F. Wyzgoski, X. Li, E. F. McCord and P. L. Rinaldi, *ACS Macro Letters*, 2013, **2**, 141–145.

(155) T. Szyperski, D. C. Yeh, D. K. Sukumaran, H. N. B. Moseley and G. T. Montelione, *Proceedings of the National Academy of Sciences*, 2002, **99**, 8009–8014.

(156) J. Garbow, D. Weitekamp and A. Pines, *Chemical Physics Letters*, 1982, **93**, 504 –509.

(157) D. Uhrin, T. Liptaj and K. Kover, *Journal of Magnetic Resonance, Series A*, 1993, **101**, 41 –46.

(158) N. Brodaczewska, Z Košťálová and D Uhrín, *Journal of Biomolecular NMR*, 2018, **70**, 115 – 122.

(159) J. Sakas and N. G. A. Bell, *Faraday Discuss.*, 2019, **218**, 191–201.

(160) P. E. Hansen, H. D. Dettman and B. D. Sykes, *Journal of Magnetic Resonance (1969)*, 1985, **62**, 487 –496.

(161) J. R. Arnold and J. Fisher, *Journal of Magnetic Resonance*, 2000, **142**, 1 –10.

(162) F. B. Mallory, E. D. Luzik, C. W. Mallory and P. J. Carroll, *The Journal of Organic Chemistry*, 1992, **57**, 366–370.

(163) D. J. Russell, C. E. Hadden, G. E. Martin and K. Krishnamurthy, *Magnetic Resonance in Chemistry*, 2002, **40**, 207–210.

(164) S. Y. Kimura, T. N. Vu, Y. Komaki, M. J. Plewa and B. J. Mariñas, *Environmental Science & Technology*, 2015, **49**, 9954–9963.

(165) I. Kristiana, H. Gallard, C. Joll and J.-P. Croué, *Water Research*, 2009, **43**, 4177 –4186.

(166) G. A. Cowman and P. C. Singer, *Environmental Science & Technology*, 1996, **30**, 16–24.

(167) H. Zhai, X. Zhang, X. Zhu, J. Liu and M. Ji, *Environmental Science & Technology*, 2014, **48**, 2579–2588.

(168) I. Schuhmann, C. B. Fotso-Fondja Yao, W. Al-Zereini, H. Anke, E. Helmke and H. Laatsch, *The Journal of Antibiotics*, 2009, **62**, 453–460.

(169) M. G. Boersma, T. Y. Dinarieva, W. J. Middelhoven, W. J. H. van Berkel, J. Doran, J. Vervoort and I. M. C. M. Rietjens, *Applied and Environmental Microbiology*, 1998, **64**, 1256–1263.

Part V

APPENDIX

```
import nmrglue as ng
import numpy as np
from skimage.util import view as blocks as vablks

path = 'C:/x/x/DataFolder
instr = "1"
outstr = "10"

fidnum = 200                          # number of FIDs in ser file
points = 4194304            # number of points in the FID e.g. 4MW

inpath = path + instr
outpath = path + outstr

dic,data = ng.bruker.read(inpath, cplex = False)

set1 = [0, 1, 2...]                  # number of each transient to
                                       # include in the summed FID

FIDs = vablks(data, block shape = (points,))

ProcData = np.zeros((points,))

for i in set1:
        ProcData += FIDs[i,:]

ng.fileio.bruker.write(path + outstr + 'set1',
                       dic, ProcData, overwrite = True)
```

Figure A.1: Python code to extract all FIDs from a ser file and allows the user to choose which FIDs they would like to sum for further processing.

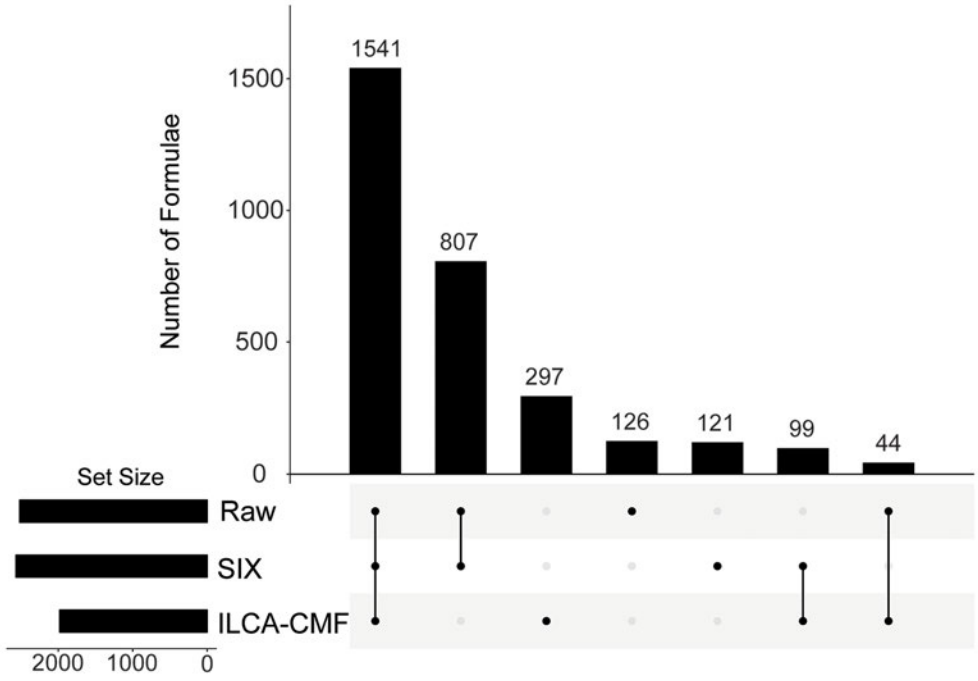


Figure A.2: UpSet plot highlighting the differences between the Raw/SIX and the ILCA-CMF samples taken in June.

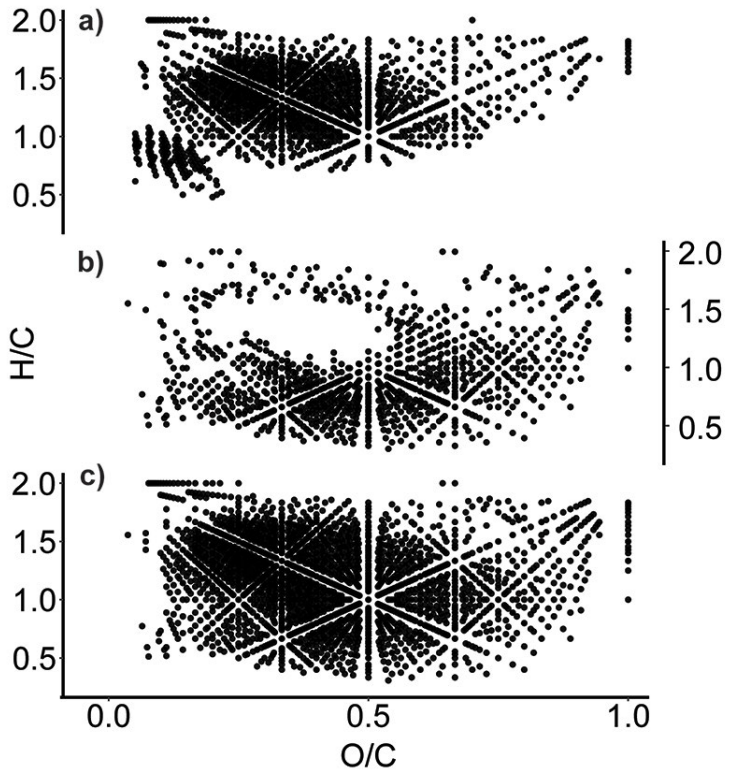


Figure A.3: Van Krevelen diagrams of **a)** the July pilot plant sample after the ILCA-CMF treatment; **b)** formulae that are no longer present after the ILCA-CMF treatment and **c)** the July SIX sample.

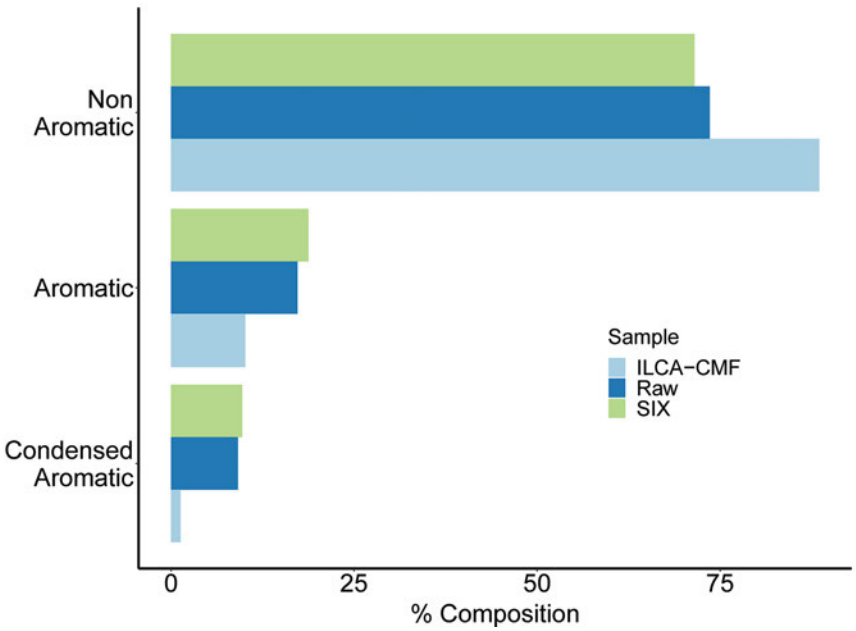


Figure A.4: AI_{mod} plot of the SIX based pilot plant June samples.

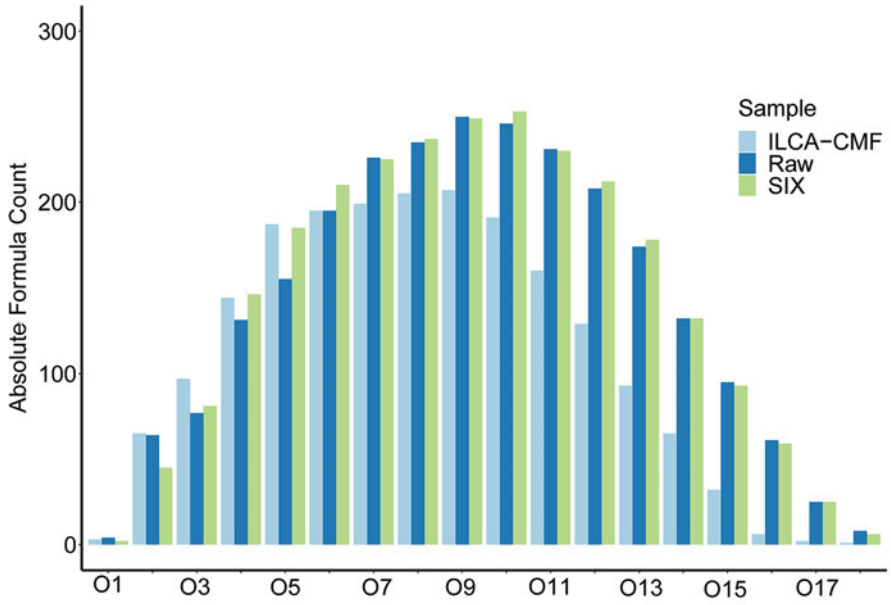


Figure A.5: Oxygen class plot of the June pilot plant samples.

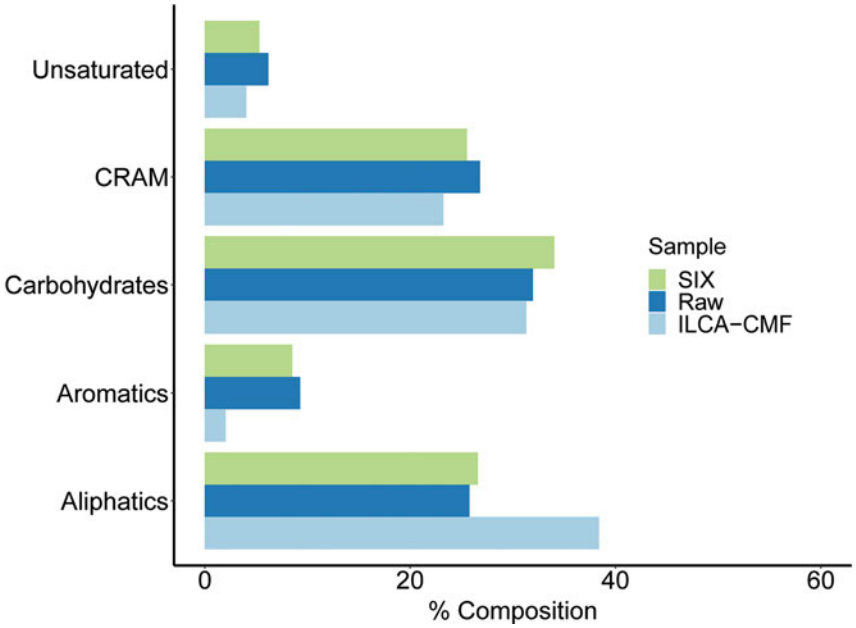


Figure A.6: NMR integration results for the June pilot plant samples.

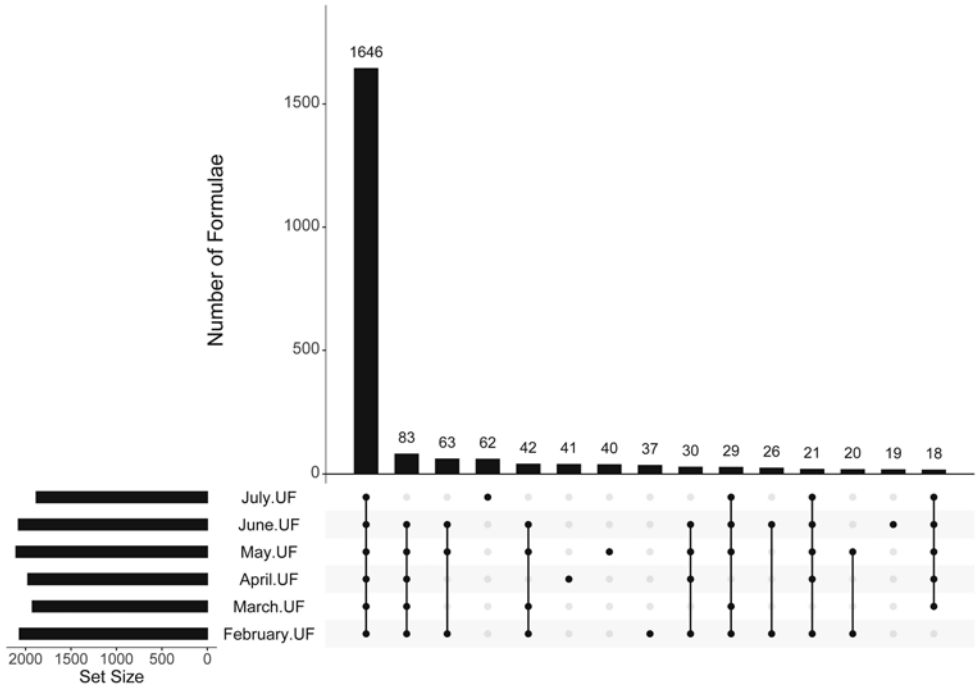


Figure B.1: UpSet plot of six UF treated water samples. Each sample represents a pair of instrumental duplicates.

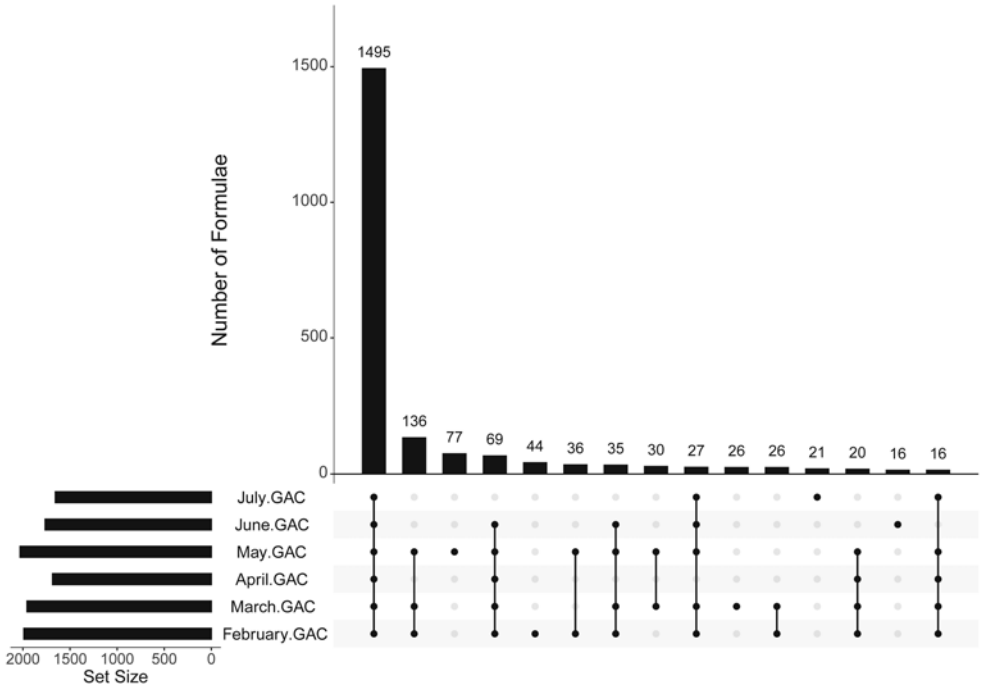


Figure B.2: UpSet plot of six GAC treatment water samples. Each sample represents a pair of instrumental duplicates.

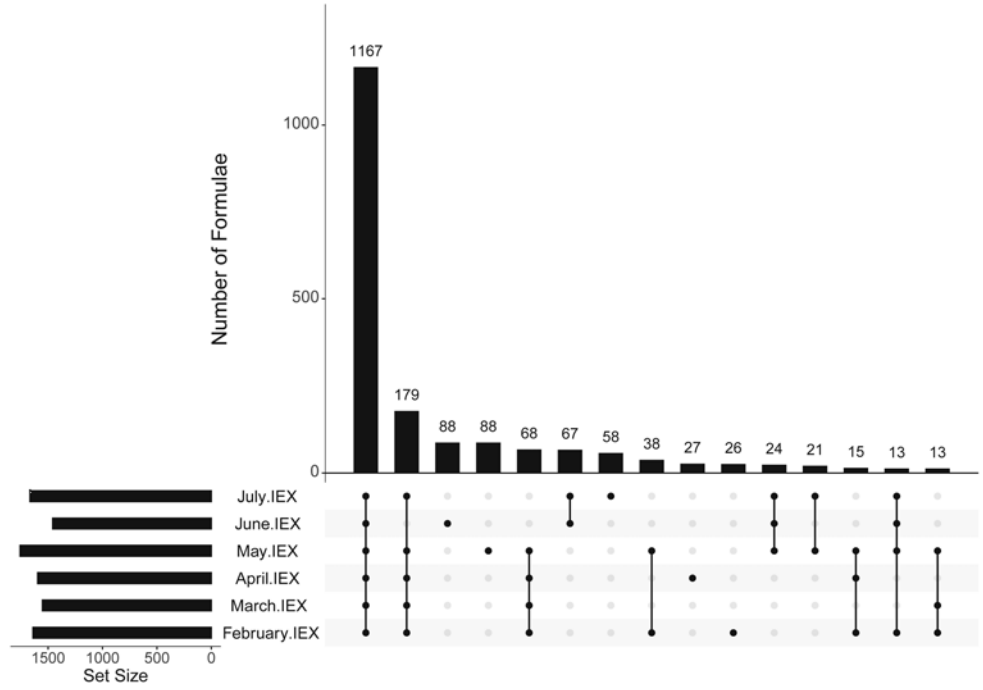


Figure B.3: UpSet plot of six IEX treated water samples. Each sample represents a pair of instrumental duplicates.

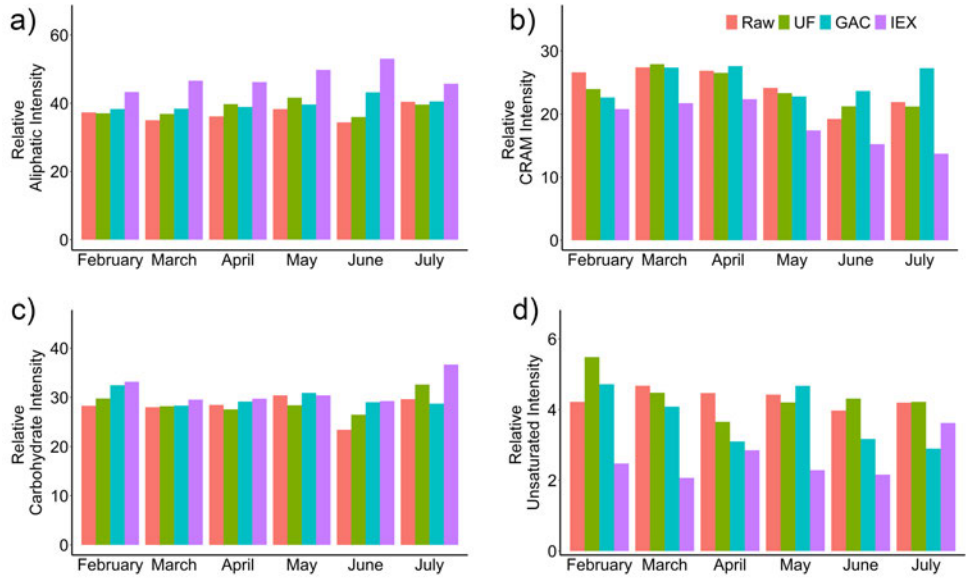


Figure B.4: Relative ^{1}H NMR intensities of the a) aliphatic region (0.5 - 2 ppm) b) CRAM region (2 - 3.1 ppm) c) carbohydrate region (3.1 - 4.6 ppm) and d) unsaturated region (4.8 - 6) for the raw, UF, GAC and IEX samples between February and July.

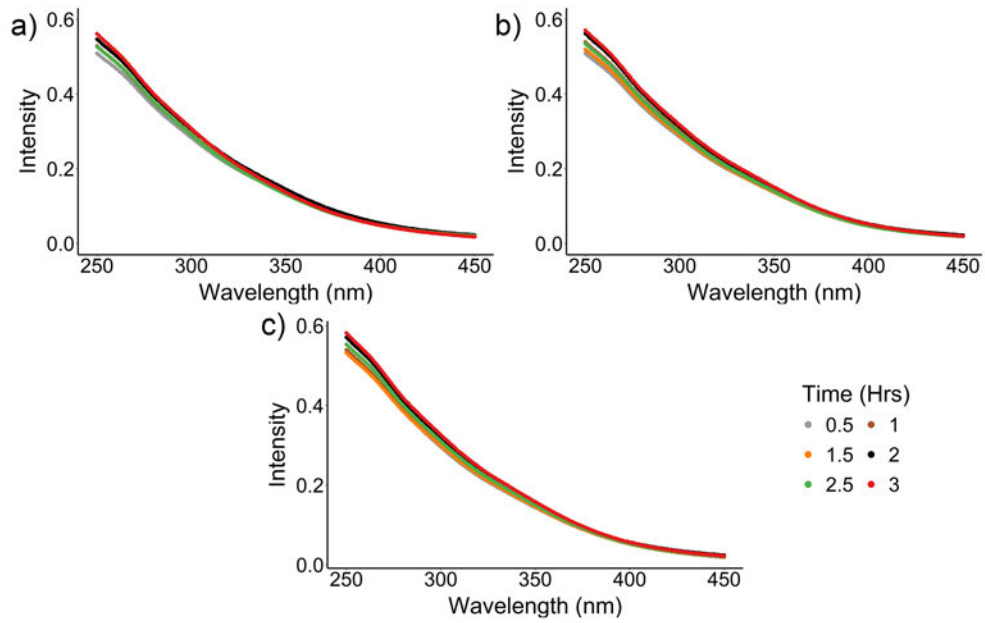


Figure C.1: UV profiles of SRFA samples after having been exposed to a) 370 nm , b) 410 nm and c) White LEDs over 3 hours.

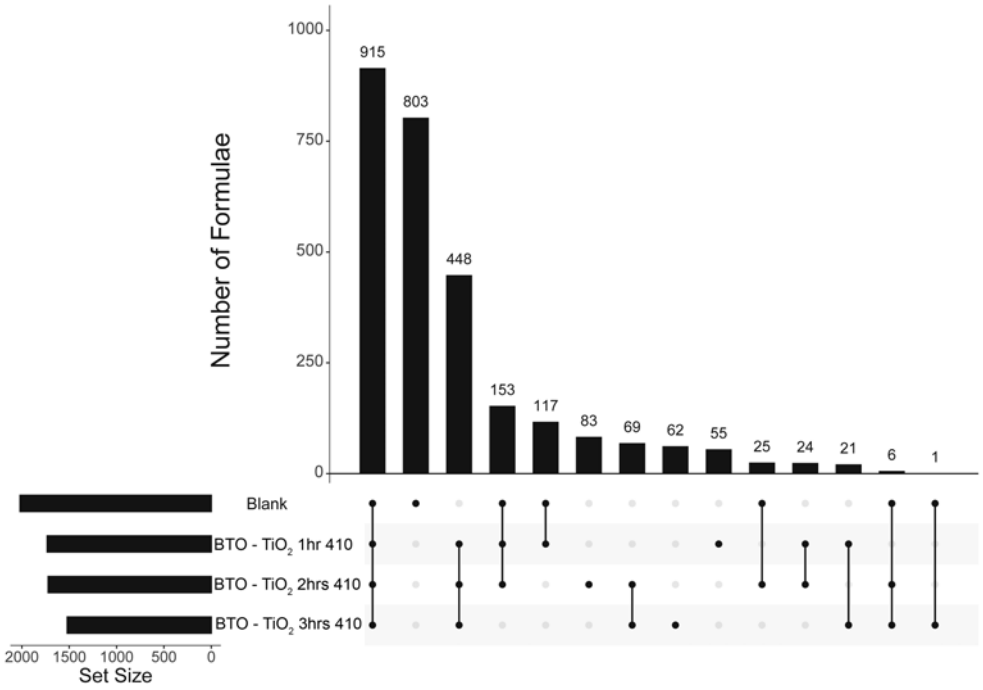


Figure C.2: UpSet plot of molecular formulae assigned in samples irradiated with the 410 nm LED for 1, 2 and 3 hours in the presence of the BTO - TiO_2 catalyst, including the reference blank data set.

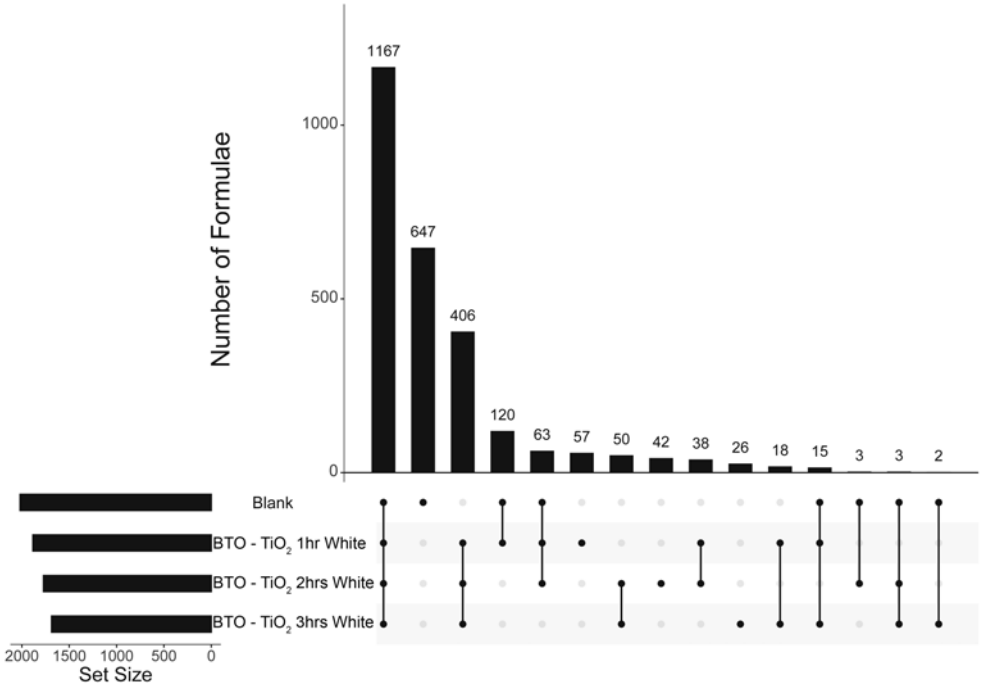


Figure C.3: UpSet plot of molecular formulae assigned in samples irradiated with the white LED for 1, 2 and 3 hours in the presence of the BTO - TiO_2 catalyst, including the reference blank data set.

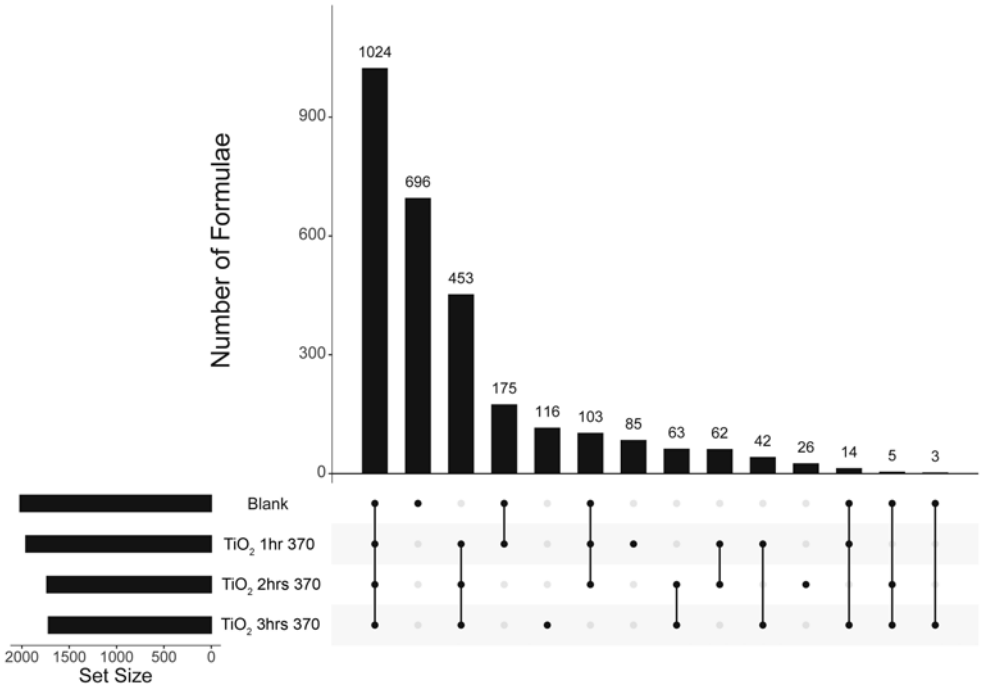


Figure C.4: UpSet plot of molecular formulae assigned in samples irradiated with 370 nm light for 1,2 and 3 hours in the presence of the pristine TiO_2 catalyst, including the reference blank data set.

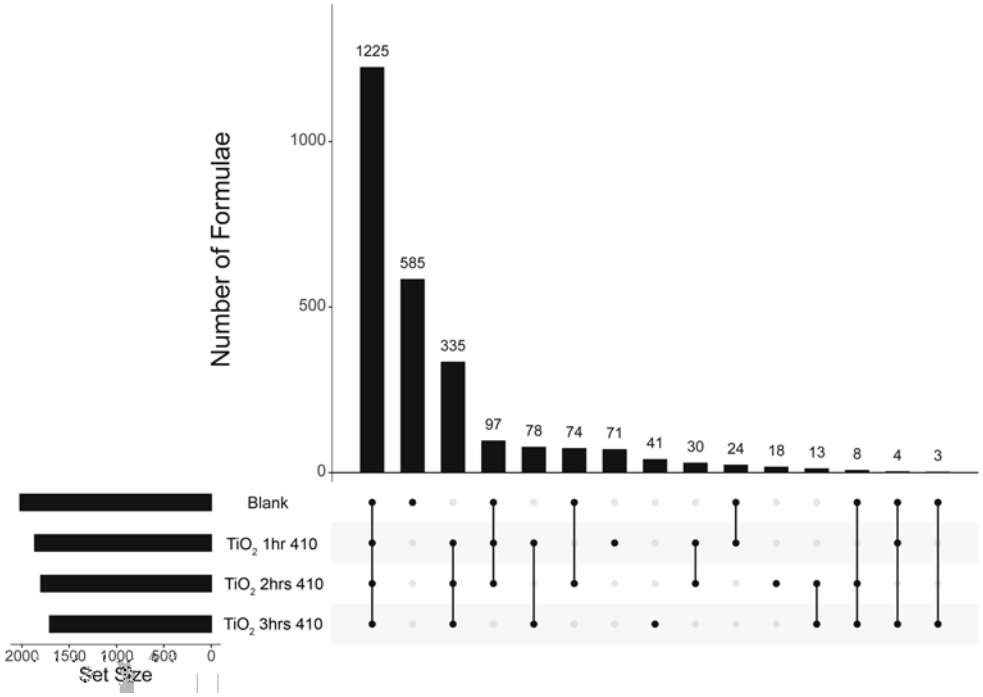


Figure C.5: UpSet plot of molecular formulae assigned in samples irradiated with 410 nm light for 1,2 and 3 hours in the presence of the pristine TiO_2 catalyst, including the reference blank data set.

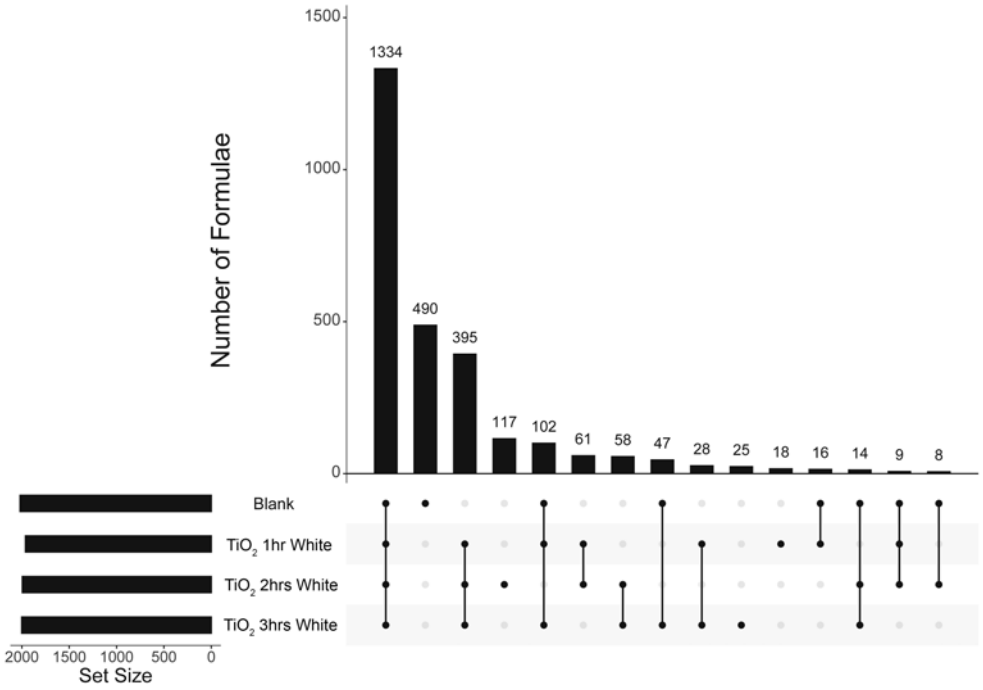


Figure C.6: UpSet plot of molecular formulae assigned in samples irradiated with the white LED for 1, 2 and 3 hours in the presence of the pristine TiO_2 catalyst, including the reference blank data set.

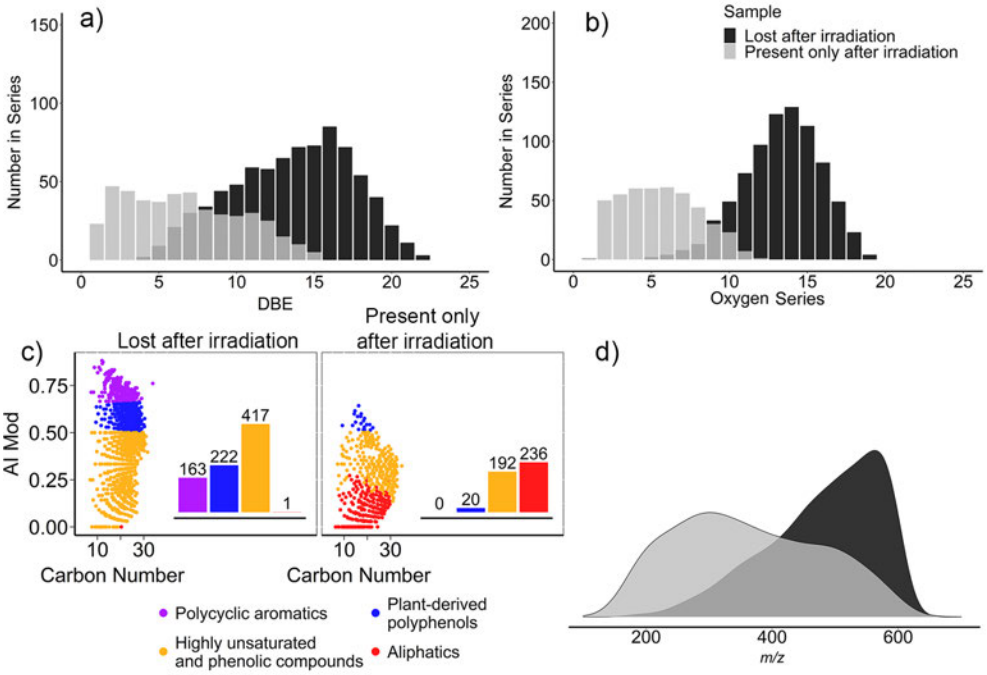


Figure C.7: Molecular characteristics of the changes to SRFA upon irradiation with a 410 nm LED in the presence of the BTO - TiO_2 catalyst. 803 degraded compounds (black) vs 448 degradation by-products (grey).

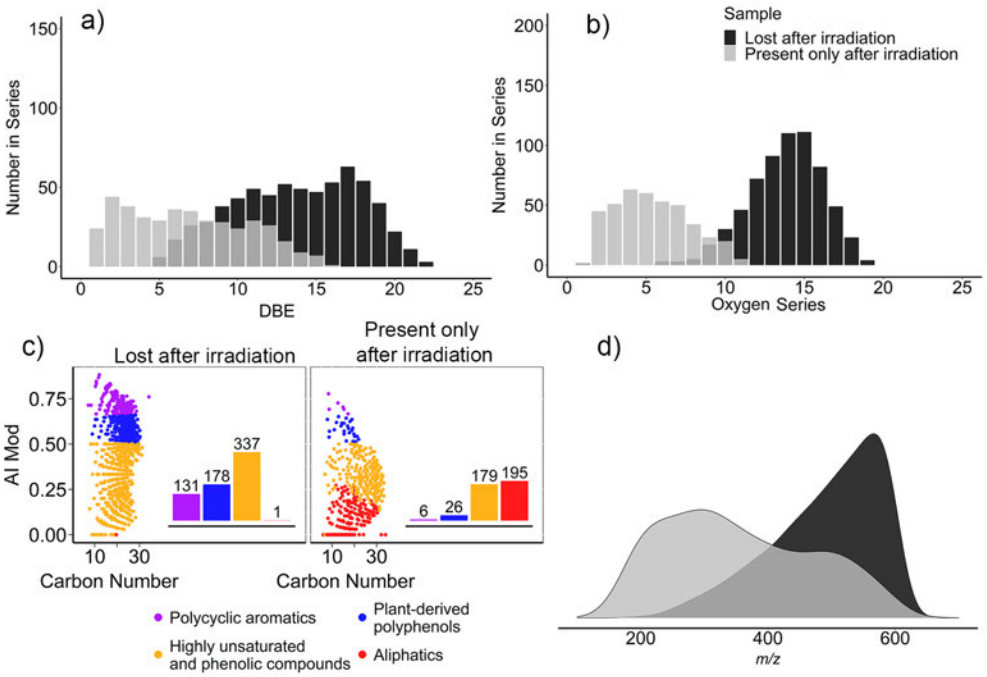


Figure C.8: Molecular characteristics of the changes to SRFA upon irradiation with a white LED array in the presence of the BTO - TiO_2 catalyst. The degraded compounds (647, black) are compared to the newly formed degradation by-products (406, grey).

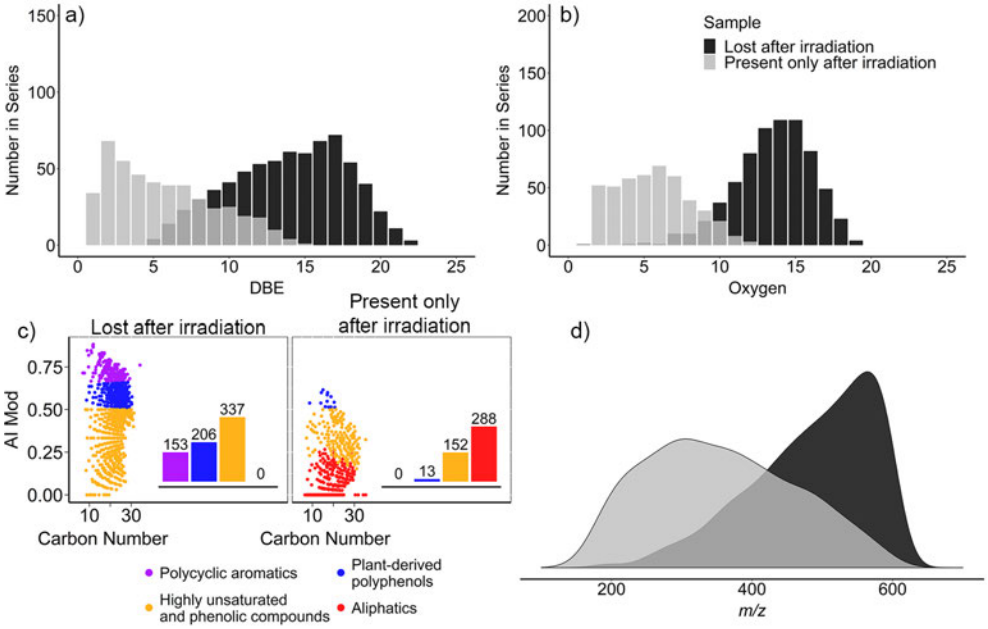


Figure C.9: Molecular characteristics of the changes to SRFA upon irradiation with a 370 nm LED in the presence of the pristine TiO_2 catalyst. The degraded compounds (696, black) are compared to the newly formed degradation by-products (453, grey).

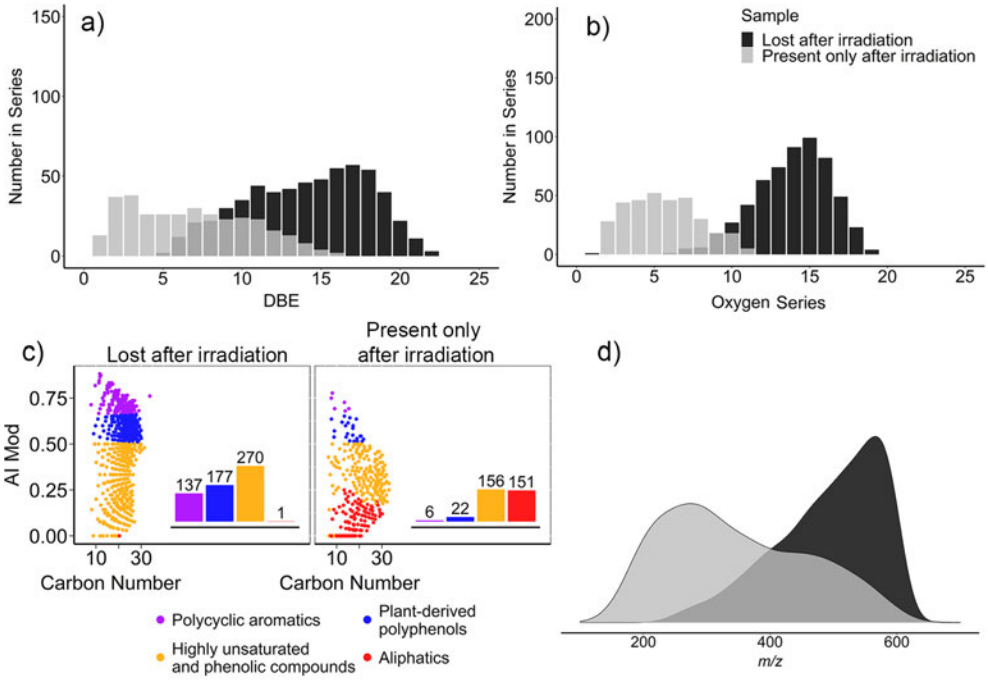


Figure C.10: Molecular characteristics of the changes to SRFA upon irradiation with a 410 nm LED in the presence of the pristine TiO_2 catalyst. The degraded compounds (585, black) are compared to the newly formed degradation by-products (335, grey).

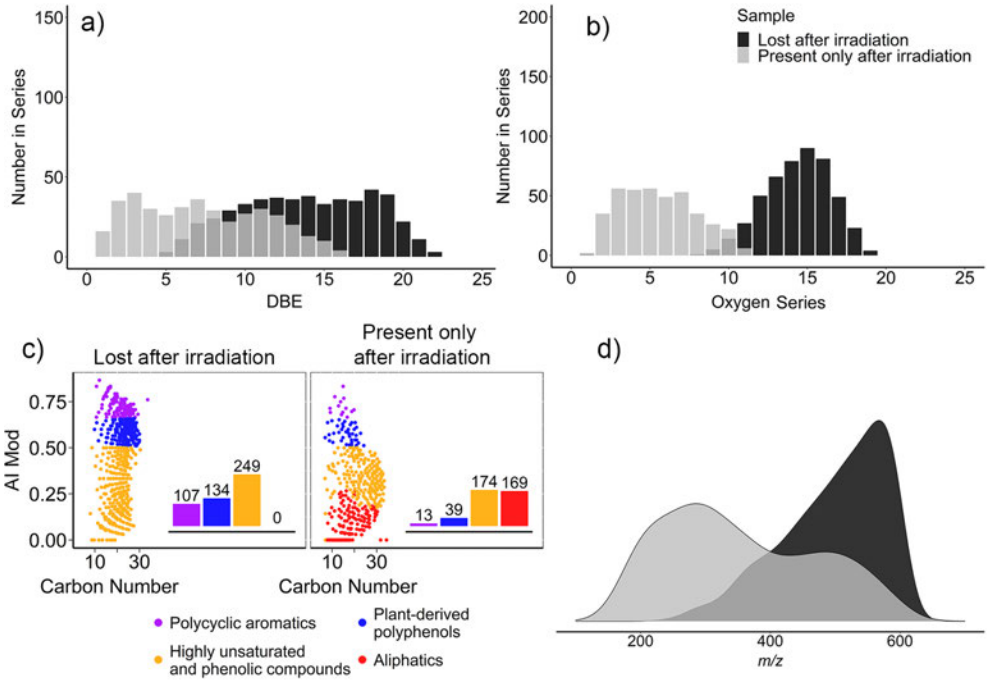


Figure C.11: Molecular characteristics of the changes to SRFA upon irradiation with the white LED array in the presence of the pristine TiO_2 catalyst. 490 degraded compounds (black) vs 395 degradation by-products (grey).

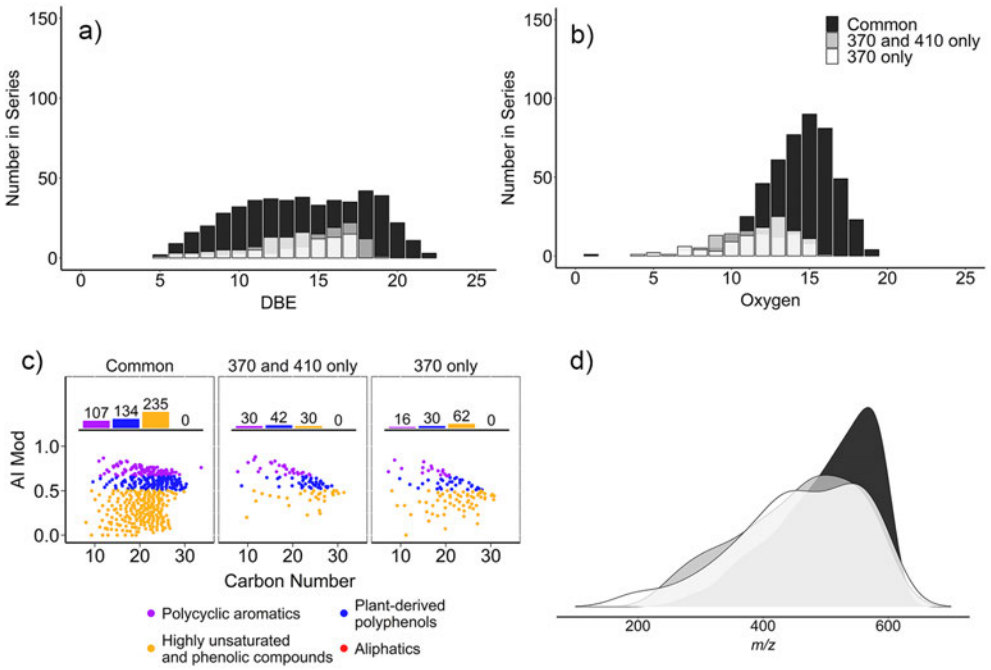


Figure C.12: Molecular characteristics of lost formulae which are common to all three wavelengths (black, 476 formulae), unique to 370 and 410 nm, (gray, 102) or unique to 370 nm (white, 108), in the presence of the TiO_2 catalyst. a) a DBE plot, b) oxygen series, c) AI_{mod} plot with a bar plot of individual compound classes and d) m/z density plot normalised independently to the sum of all points in each group.

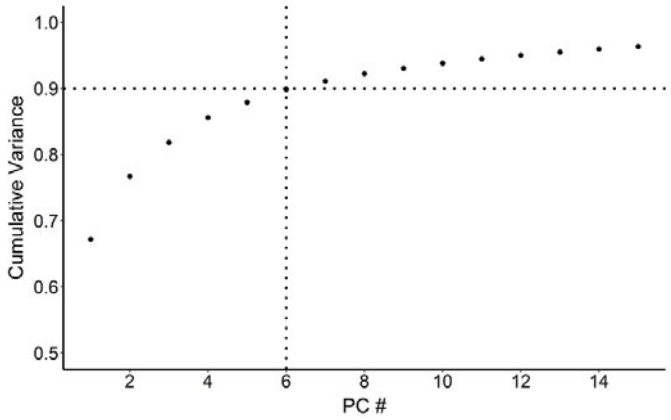


Figure C.13: Plot showing PCA principal components *vs* their cumulative variance (R^2).

APPENDIX: NMR ACQUISITION PARAMETERS AND STRUCTURE ASSIGNMENT DESCRIPTIONS

NMR spectra were acquired either on a 400 MHz Bruker Avance III nanobay system equipped with a BB TBO ^{1}H, ^{19}F probe or a 500 MHz Bruker Avance III HD spectrometer equipped with a QCI-F CryoProbe.

Table D.1: NMR experiment parameters.

Experiment	Spectrometer (MHz)	Scans	F1 (t1) Points / Sweep Width (ppm)	F2 (t2) Points / Sweep Width (ppm)	AQ time F1 / F2 (ms)
^{1}H-^{19}F COSY	400	8	768 / 10	32k / 100	96 / 433
^{1}H-^{19}F HETCOR and ^{1}H-^{19}F TOCSY-HETCOR (^{19}F detected)	500	24	320 / 10	32k / 100	32 / 350
^{1}H-^{19}F HMBC	500	32	320 / 10	32k / 100	32 / 50
1H-19F HETCOR (1H detected)	400	4	2k / 6	4k / 32	85 / 853
^{1}H-^{19}F HETCOR CP (^{1}H detected)	400	8	904 / 14	4k / 6	85 / 853
^{19}F-^{13}C ^{n}J HMBC	500	16	768 / 120	32k / 100	25 / 350
^{19}F - ^{13}C ^{1}J HMBC	500	24	768 / 120	32k / 100	25 / 350
(2, 3)D HCF	500	48	512 / 10	32k / 100	51 / 350

The remainder of this appendix presents arguments for the structural assignments of the DBPs that resulted from the chloramination of compound **3**.

3: The ^{19}F 1D spectrum shows a tall resonance at -138.87 ppm. The ^{19}F - ^{13}C HMBC indicates this ^{19}F shares a coupling (241.4 Hz) with an ipso carbon with a chemical shift of 150.85 ppm. The ^{19}F also exhibits coupling to carbons at 117.0 (ortho, 22.3 Hz), 116.98 (meta, 3.2 Hz), 126.71 (para, 2.9 Hz), 149.79 (ortho, 13.2 Hz) and 167.83 ppm (carbonyl, 2.2 Hz). The HCF experiment

indicates the ^{19}F shares couplings with protons at 6.94 (8.9 Hz), 7.66 (11.5 Hz) and 7.68 ppm (0.8 Hz). The proton - proton couplings were determined using 1D traces from the 2D CP experiment.

4: The ^{19}F 1D spectrum shows a resonance at -136.09 ppm. The ^{19}F - ^{13}C HMBC indicates this ^{19}F shares a coupling (242.5 Hz) with an ipso C at 151.30 ppm. The ^{19}F also shares couplings with carbons at 116.02 (ortho, 21.8 Hz), 118.23 (meta, 3.53 Hz), 123.23 (meta, 8.86 Hz), 124.14 (para, 3.56 Hz) and 143.93 (ortho, 12.86 Hz) ppm. The HCF experiment indicates the ^{19}F shares couplings with protons at 7.049, 6.927 and 6.860 ppm. A ortho J_{HH} was indicated but is distorted by meta coupling. Examining the size of the J_{FC} and J_{HF} couplings indicates the following combinations 116.02/7.049 and 118.23/6.860 and 124.16/6.927, confirmed by the ^{1}H - ^{13}C HSQC. The ^{1}H - ^{13}C HMBC confirmed the order of the carbons in the aromatic ring by showing long range couplings from 116.02 to 6.93, 123.25 to 6.86 and 7.049, 124.16 to 7.049, 143.93 to all protons and 151.30 to 6.86 and 7.049. Combining the data, indicates molecule **4** is 4-chloro-2-flurophenol.

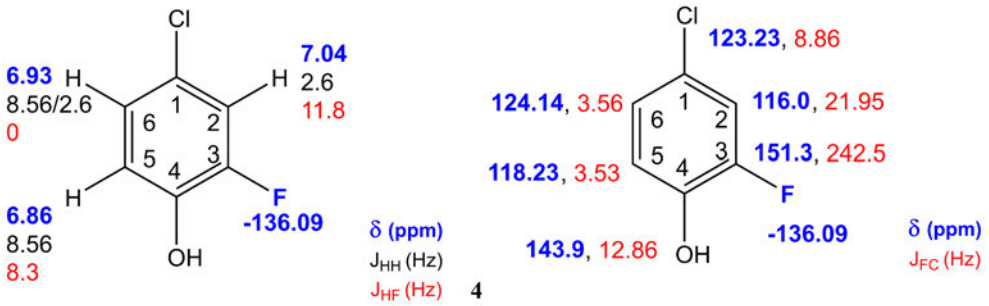
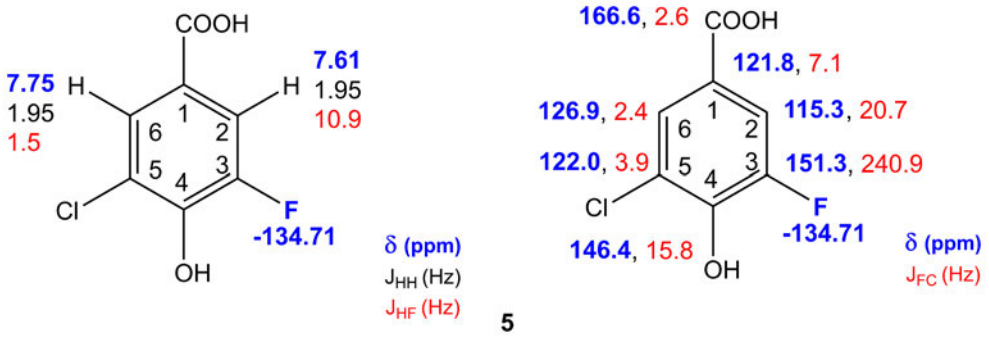


Figure D.1: Structure of compound **4**, showing chemical shifts and coupling information.

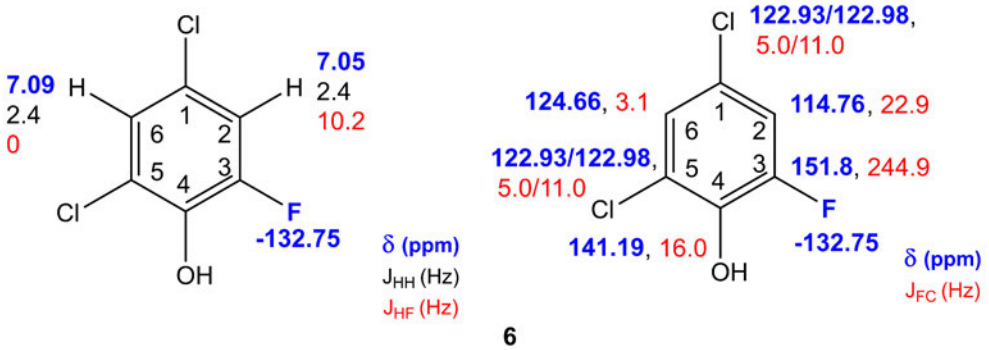
5: The ^{19}F 1D spectrum shows a tall resonance at -134.71 ppm. The ^{19}F - ^{13}C HMBC shows that the ipso carbon has a chemical shift of 151.33 ppm. The ^{19}F also exhibits coupling to the following carbons, 115.3 ppm (20.7 Hz, ortho), 146.4 ppm (15.8 Hz, ortho), 121.8 ppm (7.1 Hz, meta), 122.0 ppm (3.9 Hz, meta), 126.9 ppm (2.4 Hz, para) and 166.60 ppm (2.6 Hz, carbonyl). The HCF experiment shows the ^{19}F shares J_{HF} couplings with protons at 7.61 (10.9 Hz) and 7.75 (1.5 Hz) ppm. The J_{HH} couplings were determined using 1D traces from the 2D CP experiment. The HCF experiment confirmed which carbons were directly bonded to these protons. The ^{1}H - ^{13}C HMBC confirmed the order of the carbons in the aromatic ring. Combining the data provides compound **5**.



5

Figure D.2: Structure of compound **5** showing chemical shifts and coupling information

6: The ^{19}F 1D spectrum shows a resonance at -132.75 ppm. The ^{19}F - ^{13}C HMBC provides an ipso carbon at 151.80 ppm (244.9 Hz). The ^{19}F also shows couplings to carbons at 114.76 (23 Hz, ortho), 141.19 (16 Hz, ortho), 122.93/122.98 (5/11 Hz, meta) and 124.66 ppm (3.1 Hz, para). The HCF experiment indicates the presence of only one proton at 7.05 ppm, with a J_{HF} coupling of 10.2 directly bonded to the carbon of 114.76 ppm. Another proton, at 7.09 ppm, was only observed through the Hetcor-TOCSY experiment. J_{HH} couplings were accurately determined using 1D traces from the 2D CP experiment. The order of carbons on the aromatic ring, based on ^{19}F coupling constants, was confirmed using a ^{1}H - ^{13}C HMBC experiment, this provided structure **6**.



6

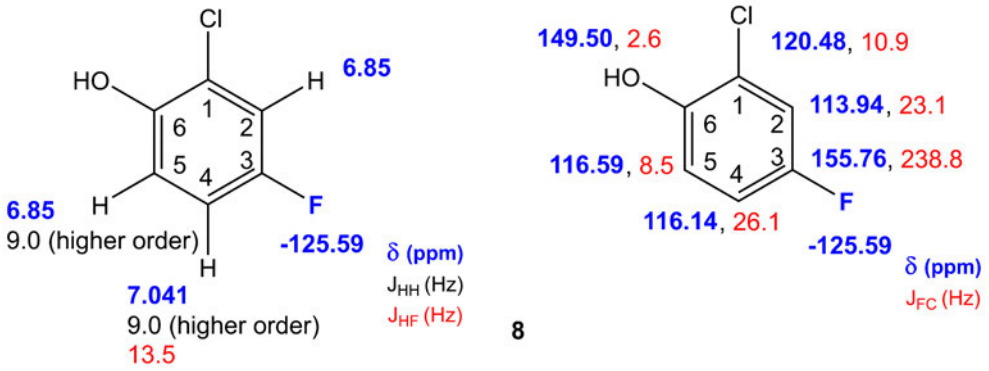
Figure D.3: Structure of compound **6** showing chemical shifts and coupling information

7: The 1D ^{19}F experiment shows a signal at -127.92 ppm, directly bonded to a carbon with a coupling constant of 234.9 Hz and chemical shift of 156.51 ppm. The ^{19}F also shares a J_{FC} coupling with two protonated carbons, the first at 115.10 (23.15 Hz, ortho) with a proton at 6.86 ppm, the second at 115.71 ppm (7.9 Hz, meta) with a proton at 6.72 ppm. Additional coupling to a quaternary carbon at 158.18 ppm (2.93 Hz, para). The data suggests that **7** is p-fluorophenol which is confirmed by the literature.[169]

7

Figure D.4: Structure of compound **7** showing chemical shifts and coupling information

8: The 1D ^{19}F experiment shows a resonance at -125.59 ppm. The ^{19}F - ^{13}C HMBC indicates this ^{19}F shares a J_{FC} of 238.8 Hz, with an ipso carbon at 155.76 ppm. There are also J_{FC} couplings to carbons with a chemical shift of 113.94 (23.1 Hz, ortho), 116.14 (26.1 Hz, ortho), 116.59 (8.5 Hz, meta), 120.48 (10.9 Hz, meta) and 149.50 (2.6 Hz, para). The HCF experiment showed the following carbon/proton combinations, 116.14 with 7.04 ppm, 113.94 with 6.85 ppm and 116.59 with 6.85 ppm. The data indicates that the structure is similar to compound **7** but with an additional chlorine ortho to the hydroxyl group, providing structure **8**.



8

Figure D.5: Structure of compound **8** showing chemical shifts and coupling information

9: The 1D ^{19}F spectra shows a ^{19}F with a chemical shift of -134.23. ^{19}F - ^{13}C HMBC experiments show carbons at 150.87 (244 Hz, ipso), 115.90 (20.9 Hz, ortho), 147.29 (16.6 Hz, ortho), 122.33 (6.6 Hz, meta), 110.66 (3.6 Hz, meta), 129.96 (2.9 Hz, para) and 166.46 (2.9 Hz, carbonyl). The HCF experiment shows that the carbons at 115.90 and 129.96 ppm are protonated, their chemical shifts are 7.64 and 7.91 ppm respectively. The data for this structure is similar to that of coumpound **5**, but the chlorine is replaced by a bromine, which is present in sodium hypochlorite.

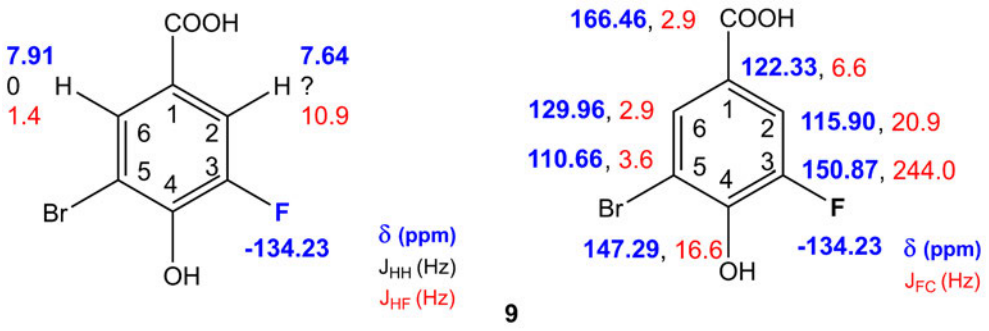


Figure D.6: Structure of compound **9** showing chemical shifts and coupling information

10: The 1D ^{19}F spectrum shows a resonance at -135.69 ppm. The ^{19}F - ^{13}C HMBC shows that this fluorine is directly coupled to a carbon at 150.23 ppm (245.4 Hz). The ^{19}F also couples to carbons at 111.86 (23.4 Hz, ortho), 151.97 (12.83 Hz, ortho), 139.59 (7.68 Hz, meta), 116.87 (3.9 Hz, meta) and a carbon at 120.82 (3.5 Hz, para). The HCF experiment shows that the carbons at 111.86, 116.87 and 120.82 ppm have protons at 7.92, 7.0 and 7.91 ppm. The ^{1}H - ^{13}C HMBC confirmed the order of carbons in the aromatic ring. The data suggests a similar substitution pattern to compound **3**. However, the protons are more deshielded, the carbon in position 1 has a much higher chemical shift and there is a lack of a carbonyl group. In the 1D ^{19}F spectrum with ^{1}H decoupling, there was an additional coupling present, 1.54 Hz. Carbon 1 also has an additional coupling of 13 Hz. This indicates the presence of ^{15}N on the carbon at position 1. Additional ^{15}N based experiments provided a chemical shift of 168.5 ppm, there was no indication of protons on this nitrogen.

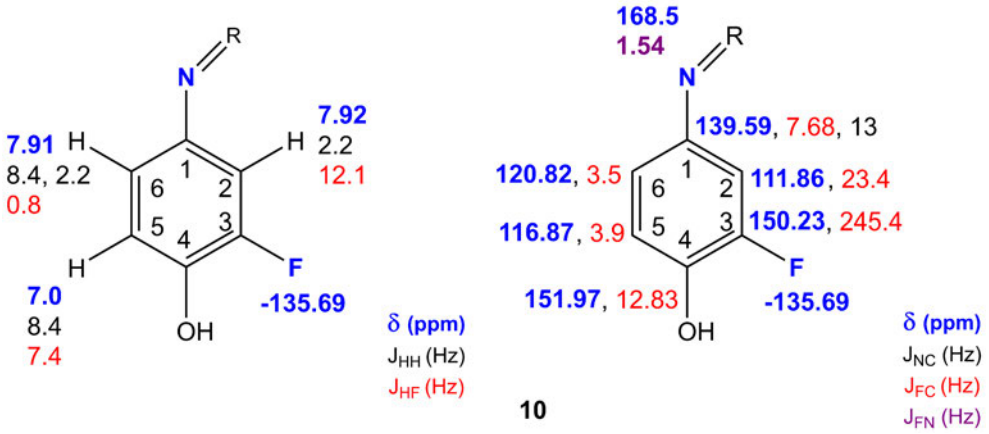


Figure D.7: Structure of compound **10** showing chemical shifts and coupling information

11: The 1D ^{19}F spectrum shows a fluorine at -132.5 ppm. The ^{19}F - ^{13}C HMBC indicates this fluorine shares a J_{FC} coupling (247.8 Hz) with an ipso carbon with a chemical shift of 152.0 ppm. The ^{19}F also shares J_{FC} couplings with carbons at 121.20 (20.2 Hz, ortho), 146.63 (16.4 Hz, ortho), 121.19 (6.7 Hz, meta), 135.96 (3.6 Hz, meta), 121.94 (3.3 Hz, para) and a carbonyl carbon at 165.50 ppm (3.6 Hz). The HCF experiment indicates that the ^{19}F couplings with two protons at 7.93 and 8.4 ppm, it also shows that these protons are directly bonded to the carbons at 121.20 and 121.94 ppm respectively. The ^{1}H - ^{13}C HMBC confirmed the position of the carbons. The data suggests that this compound has a similar substitution pattern as compound **5**, however the protons are more deshielded and carbon 5 has a higher chemical shift. Like with compound **13**, this compound was also found to have a 1.53 Hz J_{FN} coupling. There was also an additional coupling of 13 Hz on carbon 5, indicating the nitrogen is connected to this carbon. The data suggests that compound **11** is 3-fluoro-4-hydroxy-5-nitrobenzoic acid, based on the chemical shifts of 4-hydroxy-3-nitrobenzoic acid.[168]

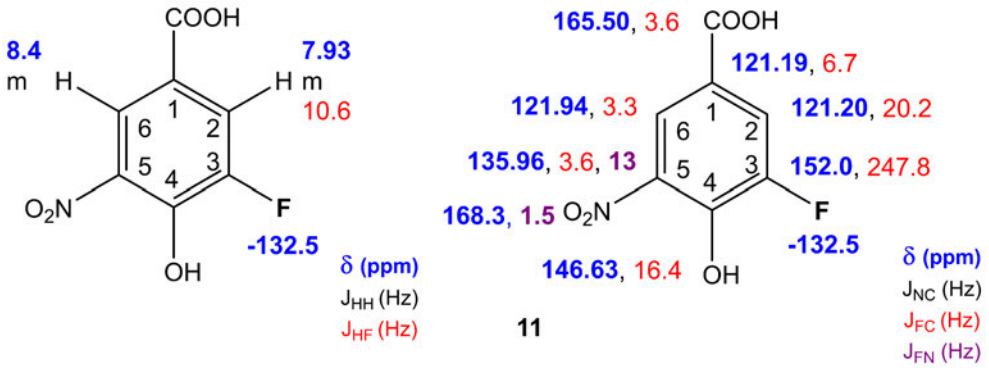


Figure D.8: Structure of compound **11** showing chemical shifts and coupling information

12: The 1D ^{19}F spectrum shows a resonance at -132.07 ppm. The ^{19}F - ^{13}C HMBC shows that the fluorine is coupled to an ipso carbon (246.8 Hz) at 150.57 ppm. The ^{19}F is also coupled to carbons at 110.51 (24.2 Hz, ortho), 148.62 (16.45 Hz, ortho), 138.72 (9.4 Hz, meta), 122.32 (5.17 Hz, meta), 121.17 (3.7 Hz, para). The HCF shows that the ^{19}F is coupled to two protons with chemical shifts of 7.9 and 8.01 ppm, it also showed that these protons are directly attached to the carbons at 110.51 and 121.17 ppm respectively. The carbon positions were confirmed with ^{1}H - ^{13}C HMBC experiments. There was an additional coupling observed on the ^{19}F of 1.68 Hz from a ^{15}N . The carbon at 138.72 ppm also exhibited a coupling of 13 Hz from this nitrogen, indicating

that the nitrogen is directly coupled to the carbon at position 1. Combining the data provides the structure 2-chloro-6-fluoro-4-nitrophenol.

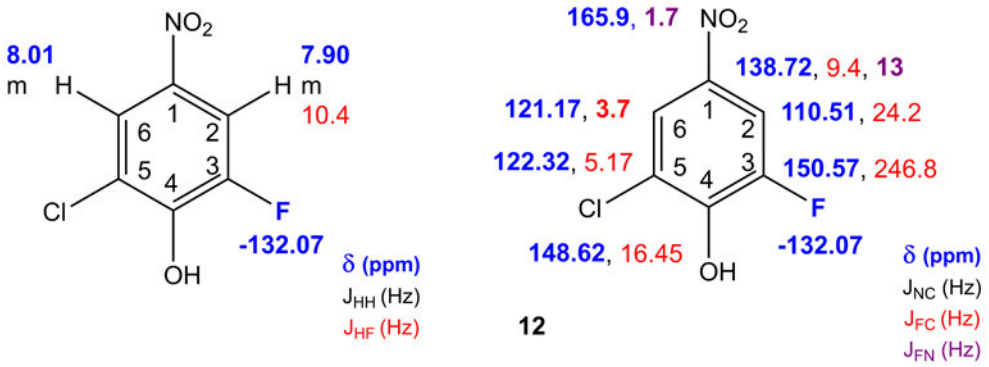


Figure D.9: Structure of compound **12** showing chemical shifts and coupling information

^{1}H - ^{19}F HETCOR EXPERIMENT - FLUORINE DETECTED

```
; hetcor_1H19F_19F-detected.as
; This experiment is fluorine detected i.e. fluorine is on channel 1
; Active 19F - 1H coupling is in antiphase, whilst the passive 19F - 1H coupling is in-phase

#include <Avance.incl>
#include <Delay.incl>
#include <Grad.incl>

"p4 = p3 * 2"
"d11 = 30m"
"d12 = 20u"
"DELTA1 = 1/4 * cnst10"
"DELTA2 = 2 * d0 + p12"
"d0 = 5u"
"in0 = inf1 / 2"
"acqt0 = -p1 * 2 / 3.1416"

1 ze
2 d11
    d1 pl2:f2 BLKGRAD
    50u UNBLKGRAD
    p1 ph1                                  ; Beginning of purge
    2u
    p1 ph2
    4u
    p16:gp1
    d16                                     ; End of purge
    (p3 ph4):f2                             ; Initial 90° pulse on 1H
    d0 pl8:f1
    (p12:sp2 ph1):f1
```

```
    d0
    DELTA1
    (center (p12:sp2 ph1) (p4 ph3):f2 )
    DELTA2
    DELTA1 pl1:f1
    (ralign (p3 ph5):f2 (p1 ph1) )
    go = 2 ph31
    d11 mc #0 to 2 F1PH(calph (ph4, +90), caldel (d0, +in0) )
exit

ph1 = 0
ph2 = 1
ph3 = 0 0 0 0 2 2 2 2
ph4 = 0 0 2 2
ph5 = 1 3
ph31 = 0 2 2 0

; pl1 : f1 channel - power level for pulse (default)
; pl2 : f2 channel - power level for pulse (default)
; pl8 : f1 channel - power level for 180 degree 19F shape pulse (1000 db)
; p1 : f1 channel - 90° high power pulse
; p3 : f2 channel - 90° high power pulse
; p4 : f2 channel - 180° high power pulse
; p12 : 180° BB 19F pulse
; sp2 : Crp140,1,20.1
; d0 : incremented delay (2D)
; d1 : relaxation delay, 1 - 5 * T1
; cnst10 = J(1H 19F) e.g. - 20 Hz
```

^{1}H - ^{19}F HETCOR EXPERIMENT - PROTON DETECTED

```
; hetcor_1H19F_1H-detected.as
; This experiment is proton detected i.e. proton is on channel 1
; Active 19F - 1H coupling is in antiphase, whilst the passive 1H - 1H coupling
is in-phase

#include <Avance.incl>
#include <Delay.incl>
#include <Grad.incl>

"d11 = 30m"
"d12 = 20u"
"d2 = 1 / 4 * cnst10"
"DELTA1 = d2 + (2 * d0) - 4u + p2"
"d0 = 5u"
"in0 = inf1 / 2"
"acqt0 = -p1 * 2 / 3.1416"

1 ze
2 d11
    d1 pl2:f2
    (p3 ph2):f2                                    ; Initial 90° pulse on 19F
    d0
    (p2 ph3)
    d0 pl8:f2
    (p12:sp2 ph1):f2
    DELTA1
    (p2 ph1)
    4u
    (p12:sp2 ph1):f2
    d2 pl2:f2
    (ralign (p1 ph1):f1 (p3 ph4):f2) )          ;INEPT transfer from 19F to 1H
    go = 2 ph31
    d11 mc #0 to 2 F1PH(calph (ph2, +90), caldel (d0, +in0) )
exit

ph1 = 0
ph2 = 0 2
```

ph3 = 0 0 0 0 2 2 2 2
ph4 = 1 1 3 3
ph31 = 0 2 2 0

; pl1 : f1 channel - power level for pulse (default)
; pl2 : f2 channel - power level for pulse (default)
; pl8 : f2 channel - power level for 180° ^{19}F shape pulse (1000 db)
; p1 : f1 channel - 90° high power pulse
; p2 : f1 channel - 180° high power pulse
; p3 : f2 channel - 90° high power pulse
; p4 : f2 channel - 180° high power pulse
; p12 : 180° BB ^{19}F pulse
; sp2 : Crp140,1,20.1
; d0 : incremented delay (2D)
; d1 : relaxation delay, 1 - 5 * T1
; cnst10 = J(^{1}H ^{19}F) e.g. - 20 Hz

^{1}H - ^{19}F TOCSY HETCOR EXPERIMENT

```
; tocsy_hetcor_1H19F.as
; This experiment is fluorine detected i.e. fluorine is on channel 1
; Magnetisation is transferred between protons using DIPSI-2 prior to INEPT
transfer to the fluorine

#include <Avance.incl>
#include <Delay.incl>
#include <Grad.incl>

"d0 = 4u"
"d11 = 30m"
"d12 = 20u"
"DELTA1 = 2 * d0 + p12"
"DELTA2 = 1 / 4 * cnst10"
"FACTOR1 = (d9 / (p6*115.112) ) / 2"
"l1 = FACTOR1 * 2"
"in0 = inf1 / 2"
"acqt0 = -p1 * 2 / 3.1416"

1 ze
2 d11
    d1 pl2:f2 pl1:f1 BLKGRAD
    50u UNBLKGRAD
    (p3 ph4):f2
    d0 pl8:f1
    (p12:sp2 ph1):f1
    d0
    (p4 ph1):f2
    DELTA1
    (p3 ph3):f2
    p16:gp1
    d16 pl10:f2

3   (p6 * 3.556 ph23):f2                    ;Beginning of DIPSI2
    (p6 * 4.556 ph25):f2
    (p6 * 3.222 ph23):f2
    (p6 * 3.167 ph25):f2
```

```
(p6 * 0.333 ph23):f2
(p6 * 2.722 ph25):f2
(p6 * 4.167 ph23):f2
(p6 * 2.944 ph25):f2
(p6 * 4.111 ph23):f2

(p6 * 3.556 ph25):f2
(p6 * 4.556 ph23):f2
(p6 * 3.222 ph25):f2
(p6 * 3.167 ph23):f2
(p6 * 0.333 ph25):f2
(p6 * 2.722 ph23):f2
(p6 * 4.167 ph25):f2
(p6 * 2.944 ph23):f2
(p6 * 4.111 ph25):f2

(p6 * 3.556 ph25):f2
(p6 * 4.556 ph23):f2
(p6 * 3.222 ph25):f2
(p6 * 3.167 ph23):f2
(p6 * 0.333 ph25):f2
(p6 * 2.722 ph23):f2
(p6 * 4.167 ph25):f2
(p6 * 2.944 ph23):f2
(p6 * 4.111 ph25):f2

(p6 * 3.556 ph23):f2
(p6 * 4.556 ph25):f2
(p6 * 3.222 ph23):f2
(p6 * 3.167 ph25):f2
(p6 * 0.333 ph23):f2
(p6 * 2.722 ph25):f2
(p6 * 4.167 ph23):f2
(p6 * 2.944 ph25):f2
(p6 * 4.111 ph23):f2
lo to 3 times l1                                    ;end of DIPSI2

10u gron0 pl0:f2 pl1:f1
(p32:sp29 ph4):f2
```

```
    20u groff
    p1 ph1
    2u
    p1 ph2
    p16:gp2
    d16 pl2:f2 pl8:f1
    (p3 ph1):f2
    DELTA2
    (center (p12:sp2 ph1) (p4 ph1):f2 )
    DELTA2 pl1:f1
    (p3 ph5):f2
    p1 ph1
    go = 2 ph31
    d11 mc #0 to 2 F1PH(calph (ph4, +90), caldel (d0, +in0) )
    d12 BLKGRAD
exit

ph1 = 0
ph2 = 1
ph3 = 0 0 0 0 2 2 2 2
ph4 = 0 0 2 2
ph5 = 1 3
ph23 = 3
ph25 = 1
ph31 = 0 2 2 0 2 0 0 2

; pl0 : f2 channel - power level for 1H shape pulse
; pl1 : f1 channel - power level for pulse (default)
; pl2 : f2 channel - power level for pulse (default)
; pl8 : f1 channel - power level for 180° 19F shape pulse (1000db)
; pl10 : f2 channel - power level for 1H TOCSY (DIPSI-2)
; p1 : f1 channel - 90° high power pulse
; p3 : f2 channel - 90° high power pulse
; p4 : f2 channel - 180° high power pulse
; p6 : f2 channel - 90° TOCSY pulse
; p12 : f1 channel - 180° BB 19F pulse
; p16 : homospoil/gradient pulse (1ms)
; p32 : z-filter (20ms)
; sp2 : Crp140,1,20.1
```

; sp29 : Crp60,20,20.10
; d9 : TOCSY mixing time
; cnst10 : J (^{19}F ^{1}H) e.g. (20 Hz)

^{1}H - ^{19}F HETCOR 1D CP EXPERIMENT

```
; hetcor_CP_1D.as
; This experiment is proton detected i.e. proton is on channel 1
; Magnetisation is transferred from fluorine using a cross polarisation flopsy-16
cycle, before allowing magnetisation to transfer amongst the protons using a
DIPSI-2 TOCSY cycle.

#include <Avance.incl>
#include <Delay.incl>
#include <Grad.incl>

"d0 = 5u"
"d11 = 30m"
"d12 = 20u"
"DELTA1 = 1 / 4 * cnst10"
"DELTA2 = (2 * d0 + p12)"
"FACTOR1 = (d9 / (p25 * 188.488) )"
"l1 = FACTOR1"
"FACTOR2 = ((d10 / (p25 * 115.112) ) / 2)"
"l2 = FACTOR2 * 2"
"in0 = inf1 / 2"
"acqt0 = -p1 * 2 / 3.1416"

1 ze
2 d11
    d12 do:f2
    d1 pl2:f2 BLKGRAD
    50u UNBLKGRAD
    (p3 ph1):f2
    p16:gp3
    d16
    4u
    p16:gp3
    d16 pl10:f1 pl11:f2

3   (p25 * 0.511 ph11):f1 (p25 * 0.511 ph11):f2      ;Beginning of FLOPSY-16
    (p25 * 1.067 ph12):f1 (p25 * 1.067 ph12):f2
    (p25 * 1.822 ph13):f1 (p25 * 1.822 ph13):f2
```

(p25 * 1.767 ph14):f1 (p25 * 1.767 ph14):f2
(p25 * 1.444 ph15):f1 (p25 * 1.444 ph15):f2
(p25 * 1.767 ph14):f1 (p25 * 1.767 ph14):f2
(p25 * 1.822 ph13):f1 (p25 * 1.822 ph13):f2
(p25 * 1.067 ph12):f1 (p25 * 1.067 ph12):f2
(p25 * 0.511 ph11):f1 (p25 * 0.511 ph11):f2

(p25 * 0.511 ph21):f1 (p25 * 0.511 ph21):f2
(p25 * 1.067 ph22):f1 (p25 * 1.067 ph22):f2
(p25 * 1.822 ph23):f1 (p25 * 1.822 ph23):f2
(p25 * 1.767 ph24):f1 (p25 * 1.767 ph24):f2
(p25 * 1.444 ph25):f1 (p25 * 1.444 ph25):f2
(p25 * 1.767 ph24):f1 (p25 * 1.767 ph24):f2
(p25 * 1.822 ph23):f1 (p25 * 1.822 ph23):f2
(p25 * 1.067 ph22):f1 (p25 * 1.067 ph22):f2
(p25 * 0.511 ph21):f1 (p25 * 0.511 ph21):f2

(p25 * 0.511 ph21):f1 (p25 * 0.511 ph21):f2
(p25 * 1.067 ph22):f1 (p25 * 1.067 ph22):f2
(p25 * 1.822 ph23):f1 (p25 * 1.822 ph23):f2
(p25 * 1.767 ph24):f1 (p25 * 1.767 ph24):f2
(p25 * 1.444 ph25):f1 (p25 * 1.444 ph25):f2
(p25 * 1.767 ph24):f1 (p25 * 1.767 ph24):f2
(p25 * 1.822 ph23):f1 (p25 * 1.822 ph23):f2
(p25 * 1.067 ph22):f1 (p25 * 1.067 ph22):f2
(p25 * 0.511 ph21):f1 (p25 * 0.511 ph21):f2

(p25 * 0.511 ph11):f1 (p25 * 0.511 ph11):f2
(p25 * 1.067 ph12):f1 (p25 * 1.067 ph12):f2
(p25 * 1.822 ph13):f1 (p25 * 1.822 ph13):f2
(p25 * 1.767 ph14):f1 (p25 * 1.767 ph14):f2
(p25 * 1.444 ph15):f1 (p25 * 1.444 ph15):f2
(p25 * 1.767 ph14):f1 (p25 * 1.767 ph14):f2
(p25 * 1.822 ph13):f1 (p25 * 1.822 ph13):f2
(p25 * 1.067 ph12):f1 (p25 * 1.067 ph12):f2
(p25 * 0.511 ph11):f1 (p25 * 0.511 ph11):f2

(p25 * 0.511 ph11):f1 (p25 * 0.511 ph11):f2
(p25 * 1.067 ph12):f1 (p25 * 1.067 ph12):f2

(p25 * 1.822 ph13):f1 (p25 * 1.822 ph13):f2
(p25 * 1.767 ph14):f1 (p25 * 1.767 ph14):f2
(p25 * 1.444 ph15):f1 (p25 * 1.444 ph15):f2
(p25 * 1.767 ph14):f1 (p25 * 1.767 ph14):f2
(p25 * 1.822 ph13):f1 (p25 * 1.822 ph13):f2
(p25 * 1.067 ph12):f1 (p25 * 1.067 ph12):f2
(p25 * 0.511 ph11):f1 (p25 * 0.511 ph11):f2

(p25 * 0.511 ph11):f1 (p25 * 0.511 ph11):f2
(p25 * 1.067 ph12):f1 (p25 * 1.067 ph12):f2
(p25 * 1.822 ph13):f1 (p25 * 1.822 ph13):f2
(p25 * 1.767 ph14):f1 (p25 * 1.767 ph14):f2
(p25 * 1.444 ph15):f1 (p25 * 1.444 ph15):f2
(p25 * 1.767 ph14):f1 (p25 * 1.767 ph14):f2
(p25 * 1.822 ph13):f1 (p25 * 1.822 ph13):f2
(p25 * 1.067 ph12):f1 (p25 * 1.067 ph12):f2
(p25 * 0.511 ph11):f1 (p25 * 0.511 ph11):f2

(p25 * 0.511 ph21):f1 (p25 * 0.511 ph21):f2
(p25 * 1.067 ph22):f1 (p25 * 1.067 ph22):f2
(p25 * 1.822 ph23):f1 (p25 * 1.822 ph23):f2
(p25 * 1.767 ph24):f1 (p25 * 1.767 ph24):f2
(p25 * 1.444 ph25):f1 (p25 * 1.444 ph25):f2
(p25 * 1.767 ph24):f1 (p25 * 1.767 ph24):f2
(p25 * 1.822 ph23):f1 (p25 * 1.822 ph23):f2
(p25 * 1.067 ph22):f1 (p25 * 1.067 ph22):f2
(p25 * 0.511 ph21):f1 (p25 * 0.511 ph21):f2

(p25 * 0.511 ph21):f1 (p25 * 0.511 ph21):f2
(p25 * 1.067 ph22):f1 (p25 * 1.067 ph22):f2
(p25 * 1.822 ph23):f1 (p25 * 1.822 ph23):f2
(p25 * 1.767 ph24):f1 (p25 * 1.767 ph24):f2
(p25 * 1.444 ph25):f1 (p25 * 1.444 ph25):f2
(p25 * 1.767 ph24):f1 (p25 * 1.767 ph24):f2
(p25 * 1.822 ph23):f1 (p25 * 1.822 ph23):f2
(p25 * 1.067 ph22):f1 (p25 * 1.067 ph22):f2
(p25 * 0.511 ph21):f1 (p25 * 0.511 ph21):f2

(p25 * 0.511 ph21):f1 (p25 * 0.511 ph21):f2

(p25 * 1.067 ph22):f1 (p25 * 1.067 ph22):f2
(p25 * 1.822 ph23):f1 (p25 * 1.822 ph23):f2
(p25 * 1.767 ph24):f1 (p25 * 1.767 ph24):f2
(p25 * 1.444 ph25):f1 (p25 * 1.444 ph25):f2
(p25 * 1.767 ph24):f1 (p25 * 1.767 ph24):f2
(p25 * 1.822 ph23):f1 (p25 * 1.822 ph23):f2
(p25 * 1.067 ph22):f1 (p25 * 1.067 ph22):f2
(p25 * 0.511 ph21):f1 (p25 * 0.511 ph21):f2

(p25 * 0.511 ph11):f1 (p25 * 0.511 ph11):f2
(p25 * 1.067 ph12):f1 (p25 * 1.067 ph12):f2
(p25 * 1.822 ph13):f1 (p25 * 1.822 ph13):f2
(p25 * 1.767 ph14):f1 (p25 * 1.767 ph14):f2
(p25 * 1.444 ph15):f1 (p25 * 1.444 ph15):f2
(p25 * 1.767 ph14):f1 (p25 * 1.767 ph14):f2
(p25 * 1.822 ph13):f1 (p25 * 1.822 ph13):f2
(p25 * 1.067 ph12):f1 (p25 * 1.067 ph12):f2
(p25 * 0.511 ph11):f1 (p25 * 0.511 ph11):f2

(p25 * 0.511 ph11):f1 (p25 * 0.511 ph11):f2
(p25 * 1.067 ph12):f1 (p25 * 1.067 ph12):f2
(p25 * 1.822 ph13):f1 (p25 * 1.822 ph13):f2
(p25 * 1.767 ph14):f1 (p25 * 1.767 ph14):f2
(p25 * 1.444 ph15):f1 (p25 * 1.444 ph15):f2
(p25 * 1.767 ph14):f1 (p25 * 1.767 ph14):f2
(p25 * 1.822 ph13):f1 (p25 * 1.822 ph13):f2
(p25 * 1.067 ph12):f1 (p25 * 1.067 ph12):f2
(p25 * 0.511 ph11):f1 (p25 * 0.511 ph11):f2

(p25 * 0.511 ph21):f1 (p25 * 0.511 ph21):f2
(p25 * 1.067 ph22):f1 (p25 * 1.067 ph22):f2
(p25 * 1.822 ph23):f1 (p25 * 1.822 ph23):f2
(p25 * 1.767 ph24):f1 (p25 * 1.767 ph24):f2
(p25 * 1.444 ph25):f1 (p25 * 1.444 ph25):f2
(p25 * 1.767 ph24):f1 (p25 * 1.767 ph24):f2
(p25 * 1.822 ph23):f1 (p25 * 1.822 ph23):f2
(p25 * 1.067 ph22):f1 (p25 * 1.067 ph22):f2
(p25 * 0.511 ph21):f1 (p25 * 0.511 ph21):f2

(p25 * 0.511 ph21):f1 (p25 * 0.511 ph21):f2
(p25 * 1.067 ph22):f1 (p25 * 1.067 ph22):f2
(p25 * 1.822 ph23):f1 (p25 * 1.822 ph23):f2
(p25 * 1.767 ph24):f1 (p25 * 1.767 ph24):f2
(p25 * 1.444 ph25):f1 (p25 * 1.444 ph25):f2
(p25 * 1.767 ph24):f1 (p25 * 1.767 ph24):f2
(p25 * 1.822 ph23):f1 (p25 * 1.822 ph23):f2
(p25 * 1.067 ph22):f1 (p25 * 1.067 ph22):f2
(p25 * 0.511 ph21):f1 (p25 * 0.511 ph21):f2

(p25 * 0.511 ph21):f1 (p25 * 0.511 ph21):f2
(p25 * 1.067 ph22):f1 (p25 * 1.067 ph22):f2
(p25 * 1.822 ph23):f1 (p25 * 1.822 ph23):f2
(p25 * 1.767 ph24):f1 (p25 * 1.767 ph24):f2
(p25 * 1.444 ph25):f1 (p25 * 1.444 ph25):f2
(p25 * 1.767 ph24):f1 (p25 * 1.767 ph24):f2
(p25 * 1.822 ph23):f1 (p25 * 1.822 ph23):f2
(p25 * 1.067 ph22):f1 (p25 * 1.067 ph22):f2
(p25 * 0.511 ph21):f1 (p25 * 0.511 ph21):f2

(p25 * 0.511 ph11):f1 (p25 * 0.511 ph11):f2
(p25 * 1.067 ph12):f1 (p25 * 1.067 ph12):f2
(p25 * 1.822 ph13):f1 (p25 * 1.822 ph13):f2
(p25 * 1.767 ph14):f1 (p25 * 1.767 ph14):f2
(p25 * 1.444 ph15):f1 (p25 * 1.444 ph15):f2
(p25 * 1.767 ph14):f1 (p25 * 1.767 ph14):f2
(p25 * 1.822 ph13):f1 (p25 * 1.822 ph13):f2
(p25 * 1.067 ph12):f1 (p25 * 1.067 ph12):f2
(p25 * 0.511 ph11):f1 (p25 * 0.511 ph11):f2

(p25 * 0.511 ph11):f1 (p25 * 0.511 ph11):f2
(p25 * 1.067 ph12):f1 (p25 * 1.067 ph12):f2
(p25 * 1.822 ph13):f1 (p25 * 1.822 ph13):f2
(p25 * 1.767 ph14):f1 (p25 * 1.767 ph14):f2
(p25 * 1.444 ph15):f1 (p25 * 1.444 ph15):f2
(p25 * 1.767 ph14):f1 (p25 * 1.767 ph14):f2
(p25 * 1.822 ph13):f1 (p25 * 1.822 ph13):f2
(p25 * 1.067 ph12):f1 (p25 * 1.067 ph12):f2
(p25 * 0.511 ph11):f1 (p25 * 0.511 ph11):f2

```
    lo to 3 times l1                                    ;End of FLOPSY-16

    (p8 ph4):f1
    p25 ph3
    10u gron0
    (p32:sp29 ph4):f1
    20u groff
    d16 pl10:f1

4   p6 * 3.556 ph26                                     ;Beginning of DIPSI-2
    p6 * 4.556 ph27
    p6 * 3.222 ph26
    p6 * 3.167 ph27
    p6 * 0.333 ph26
    p6 * 2.722 ph27
    p6 * 4.167 ph26
    p6 * 2.944 ph27
    p6 * 4.111 ph26

    p6 * 3.556 ph27
    p6 * 4.556 ph26
    p6 * 3.222 ph27
    p6 * 3.167 ph26
    p6 * 0.333 ph27
    p6 * 2.722 ph26
    p6 * 4.167 ph27
    p6 * 2.944 ph26
    p6 * 4.111 ph27

    p6 * 3.556 ph27
    p6 * 4.556 ph26
    p6 * 3.222 ph27
    p6 * 3.167 ph26
    p6 * 0.333 ph27
    p6 * 2.722 ph26
    p6 * 4.167 ph27
    p6 * 2.944 ph26
    p6 * 4.111 ph27
```

```
    p6 * 3.556 ph26
    p6 * 4.556 ph27
    p6 * 3.222 ph26
    p6 * 3.167 ph27
    p6 * 0.333 ph26
    p6 * 2.722 ph27
    p6 * 4.167 ph26
    p6 * 2.944 ph27
    p6 * 4.111 ph26
    lo to 4 times l2                                    ;End of DIPSI-2

    p16:gp1
    d16
    10u gron0 * 1.333
    (p32 * 0.75:sp29 ph4):f1
    20u groff
    d16 pl1:f1 pl16:f2
    p1 ph5
    go = 2 ph31 cpd3:f2
    30m do:f2 mc #o to 2 F0 (zd)
    4u BLKGRAD
exit

ph1 = 1 3
ph2 = 0 0 1 1
ph3 = 1 1 1 1 3 3 3 3
ph4 = 0
ph5 = 2
ph11 = (720) 0
ph12 = (720) 90
ph13 = (720) 135
ph14 = (720) 630
ph15 = (720) 45
ph21 = (720) 360
ph22 = (720) 450
ph23 = (720) 495
ph24 = (720) 270
ph25 = (720) 405
ph26 = 3
```

ph27 = 1
ph31 = 0 2 2 0 2 0 0 2

; pl1 : f1 channel - power level for pulse (default)
; pl2 : f2 channel - power level for pulse (default)
; pl10 : f1 channel - power level for FLOPSY-16 and DIPSI-2
; pl11 : f2 channel - power level for FLOPSY-16
; pl16 : f2 channel - power level for decoupling
; p6 : f1 channel - TOCSY pulse
; p12 : 180° selective pulse
; p25 : pulse for FLOPSY spin lock
; p32 : z filter pulse
; sp2 : Gaus1_180r.1000
; sp29 : Crp60,20,20.10
; cpd3 : garp4
; cnst10 : J (^{19}F ^{1}H) e.g. (20 Hz)

^{1}H - ^{19}F HETCOR 2D CP EXPERIMENT

```
; hetcor_CP_2D.as
; This experiment is proton detected i.e. proton is on channel 1
; Magnetisation is transferred from fluorine using a cross polarisation DIPSI-3
cycle, before allowing magnetisation to transfer amongst the protons using a
DIPSI-2 TOCSY cycle.

#include <Avance.incl>
#include <Delay.incl>
#include <Grad.incl>

"d0 = 1u"
"d11 = 30m"
"d12 = 20u"
"DELTA1 = 1 / 4 * csnt10"
"DELTA2 = (2 * d0 + p12)
"FACTOR1 = (d9 / (p25 * 217.3) )
"l1 = FACTOR1"
"FACTOR2 = ((d10 / (p6 * 115.112) ) / 2)"
"l2 = FACTOR2 * 2"
"in0 = inf1 / 2"
"acqt0 = (-p1 * (2 / 3.1416) )"

1 ze
2 d11
    d12 do:f2
    d1 pl2:f2 pl1:f1 BLKGRAD
    50u UNBLKGRAD
    (p3 ph1):f2
    d0
    p2 ph4
    d0 pl10:f1 pl11:f2
    (p8 ph11):f2

3   (p25 * 2.722 ph11):f1 (p25 * 2.722 ph13):f2          ;Beginning of DIPSI-3
    (p25 * 4.389 ph12):f1 (p25 * 4.389 ph14):f2
    (p25 * 2.778 ph11):f1 (p25 * 2.778 ph13):f2
    (p25 * 3.056 ph12):f1 (p25 * 3.056 ph14):f2
```

(p25 * 0.333 ph11):f1 (p25 * 0.333 ph13):f2
(p25 * 2.556 ph12):f1 (p25 * 2.556 ph14):f2
(p25 * 4.000 ph11):f1 (p25 * 4.000 ph13):f2
(p25 * 2.722 ph12):f1 (p25 * 2.722 ph14):f2
(p25 * 4.111 ph11):f1 (p25 * 4.111 ph13):f2
(p25 * 3.778 ph12):f1 (p25 * 3.778 ph14):f2
(p25 * 3.889 ph11):f1 (p25 * 3.889 ph13):f2
(p25 * 2.889 ph12):f1 (p25 * 2.889 ph14):f2
(p25 * 3.000 ph11):f1 (p25 * 3.000 ph13):f2
(p25 * 0.333 ph12):f1 (p25 * 0.333 ph14):f2
(p25 * 2.500 ph11):f1 (p25 * 2.500 ph13):f2
(p25 * 4.050 ph12):f1 (p25 * 4.050 ph14):f2
(p25 * 2.830 ph11):f1 (p25 * 2.830 ph13):f2
(p25 * 4.389 ph12):f1 (p25 * 4.389 ph14):f2

(p25 * 2.722 ph12):f1 (p25 * 2.722 ph14):f2
(p25 * 4.389 ph11):f1 (p25 * 4.389 ph13):f2
(p25 * 2.778 ph12):f1 (p25 * 2.778 ph14):f2
(p25 * 3.056 ph11):f1 (p25 * 3.056 ph13):f2
(p25 * 0.333 ph12):f1 (p25 * 0.333 ph14):f2
(p25 * 2.556 ph11):f1 (p25 * 2.556 ph13):f2
(p25 * 4.000 ph12):f1 (p25 * 4.000 ph14):f2
(p25 * 2.722 ph11):f1 (p25 * 2.722 ph13):f2
(p25 * 4.111 ph12):f1 (p25 * 4.111 ph14):f2
(p25 * 3.778 ph11):f1 (p25 * 3.778 ph13):f2
(p25 * 3.889 ph12):f1 (p25 * 3.889 ph14):f2
(p25 * 2.889 ph11):f1 (p25 * 2.889 ph13):f2
(p25 * 3.000 ph12):f1 (p25 * 3.000 ph14):f2
(p25 * 0.333 ph11):f1 (p25 * 0.333 ph13):f2
(p25 * 2.500 ph12):f1 (p25 * 2.500 ph14):f2
(p25 * 4.050 ph11):f1 (p25 * 4.050 ph13):f2
(p25 * 2.830 ph12):f1 (p25 * 2.830 ph14):f2
(p25 * 4.389 ph11):f1 (p25 * 4.389 ph13):f2

(p25 * 2.722 ph12):f1 (p25 * 2.722 ph14):f2
(p25 * 4.389 ph11):f1 (p25 * 4.389 ph13):f2
(p25 * 2.778 ph12):f1 (p25 * 2.778 ph14):f2
(p25 * 3.056 ph11):f1 (p25 * 3.056 ph13):f2
(p25 * 0.333 ph12):f1 (p25 * 0.333 ph14):f2

```
(p25 * 2.556 ph11):f1 (p25 * 2.556 ph13):f2
(p25 * 4.000 ph12):f1 (p25 * 4.000 ph14):f2
(p25 * 2.722 ph11):f1 (p25 * 2.722 ph13):f2
(p25 * 4.111 ph12):f1 (p25 * 4.111 ph14):f2
(p25 * 3.778 ph11):f1 (p25 * 3.778 ph13):f2
(p25 * 3.889 ph12):f1 (p25 * 3.889 ph14):f2
(p25 * 2.889 ph11):f1 (p25 * 2.889 ph13):f2
(p25 * 3.000 ph12):f1 (p25 * 3.000 ph14):f2
(p25 * 0.333 ph11):f1 (p25 * 0.333 ph13):f2
(p25 * 2.500 ph12):f1 (p25 * 2.500 ph14):f2
(p25 * 4.050 ph11):f1 (p25 * 4.050 ph13):f2
(p25 * 2.830 ph12):f1 (p25 * 2.830 ph14):f2
(p25 * 4.389 ph11):f1 (p25 * 4.389 ph13):f2

(p25 * 2.722 ph11):f1 (p25 * 2.722 ph13):f2
(p25 * 4.389 ph12):f1 (p25 * 4.389 ph14):f2
(p25 * 2.778 ph11):f1 (p25 * 2.778 ph13):f2
(p25 * 3.056 ph12):f1 (p25 * 3.056 ph14):f2
(p25 * 0.333 ph11):f1 (p25 * 0.333 ph13):f2
(p25 * 2.556 ph12):f1 (p25 * 2.556 ph14):f2
(p25 * 4.000 ph11):f1 (p25 * 4.000 ph13):f2
(p25 * 2.722 ph12):f1 (p25 * 2.722 ph14):f2
(p25 * 4.111 ph11):f1 (p25 * 4.111 ph13):f2
(p25 * 3.778 ph12):f1 (p25 * 3.778 ph14):f2
(p25 * 3.889 ph11):f1 (p25 * 3.889 ph13):f2
(p25 * 2.889 ph12):f1 (p25 * 2.889 ph14):f2
(p25 * 3.000 ph11):f1 (p25 * 3.000 ph13):f2
(p25 * 0.333 ph12):f1 (p25 * 0.333 ph14):f2
(p25 * 2.500 ph11):f1 (p25 * 2.500 ph13):f2
(p25 * 4.050 ph12):f1 (p25 * 4.050 ph14):f2
(p25 * 2.830 ph11):f1 (p25 * 2.830 ph13):f2
(p25 * 4.389 ph12):f1 (p25 * 4.389 ph14):f2
lo to 3 times l1                                    ;End of DIPSI-3

p8 ph4
p25 ph3
10u gron0
(p32:sp29 ph4):f1
20u groff
```

```
    d16 pl10:f1

4   p6 * 3.556 ph23                          ;Beginning of DIPSI-2
    p6 * 4.556 ph25
    p6 * 3.222 ph23
    p6 * 3.167 ph25
    p6 * 0.333 ph23
    p6 * 2.722 ph25
    p6 * 4.167 ph23
    p6 * 2.944 ph25
    p6 * 4.111 ph23

    p6 * 3.556 ph25
    p6 * 4.556 ph23
    p6 * 3.222 ph25
    p6 * 3.167 ph23
    p6 * 0.333 ph25
    p6 * 2.722 ph23
    p6 * 4.167 ph25
    p6 * 2.944 ph23
    p6 * 4.111 ph25

    p6 * 3.556 ph25
    p6 * 4.556 ph23
    p6 * 3.222 ph25
    p6 * 3.167 ph23
    p6 * 0.333 ph25
    p6 * 2.722 ph23
    p6 * 4.167 ph25
    p6 * 2.944 ph23
    p6 * 4.111 ph25

    p6 * 3.556 ph23
    p6 * 4.556 ph25
    p6 * 3.222 ph23
    p6 * 3.167 ph25
    p6 * 0.333 ph23
    p6 * 2.722 ph25
    p6 * 4.167 ph23
```

```
    p6 * 2.944 ph25
    p6 * 4.111 ph23
    lo to 4 times l2                                    ;End of DIPSI-2

    p16:gp1
    d16
    10u gron0 * 1.333
    (p32 * 0.75:sp29 ph4):f1
    20u groff
    d16 pl1:f1 pl16:f2
    p1 ph6
    go = 2 ph31 cpd3:f2
    d11 do:f2 mc #0 to 2 F1PH(calph (ph1, +90), caldel (d-, +in0) )
    4u BLKGRAD
exit

ph1 = 1 3
ph3 = 1 1 3 3
ph4 = 0
ph6 = 2
ph11 = 0
ph12 = 2
ph13 = 0
ph14 = 2
ph23 = 3
ph25 = 1
ph31 = 0 2 2 0

; pl1 : f1 channel - power level for pulse (default)
; pl2 : f2 channel - power level for pulse (default)
; pl10 : f1 channel - power level for DIPSI-2 and DIPSI-3
; pl11 : f2 channel - power level for DIPSI-3
; pl16 : f2 channel - power level for decoupling
; p6 : f1 channel - TOCSY pulse
; p8 : flip back pulse (return magnetisation to z)
; p25 : pulse for spin lock
; p32 : z filter pulse
; sp29 : Crp60,20,20.10
; cpd3 : garp4
```

; cnst10 : J (^{19}F ^{1}H) e.g. (20 Hz)

^{19}F - ^{13}C ^{1}J HMBC EXPERIMENT

```
#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>
"cnst30 = (1 - sfo2 / sfo1) / (1 + sfo2 / sfo1)"
define list<gradient> EA1 =  1.000 -cnst30 define list<gradient> EA2 =
-cnst30 1.000

"p2 = p1 * 2"
"p4 = p3 * 2"
"d6 = 1s / (cnst13 * 4)"
"d11 = 30m"
"d12 = 20u"
"in0 = inf1 / 2"
"acqt0 = 0"
"d0 = 3u"

"DELTA = p1 * 2 / 3.1416"
"DELTA1 = p12 + d0 * 2"
"DELTA2 = 2 * p16 + 2 * d16 + p24 + DELTA1"
"DELTA3 = 2 * p16 + 2 * d16 + p24 + DELTA1"
"DELTA4 = (DELTA3) / 2 - d6 - (p14/2)"
"DELTA5 = d6 + (DELTA3 / 2) - (p14 / 2)"

1 ze
    d12 pl16:f3
2 d11 do:f3
    d12 BLKGRAD
    d1 pl1:f1 pl0:f2
3 d12 cpd2:f3
    p1 ph1
    20u pl9:f1
    (p12:sp2 ph2):f1
    20u
    DELTA4
    (p14:sp3 ph1):f2
```

```
    DELTA5 pl2:f2 UNBLKGRAD
    (p3 ph3):f2
    d0
    (p12:sp2 ph2):f1
    d0
    4u do:f3
    p16:gp1 * EA1
    d16 pl0:f2
    (p24:sp7 ph4):f2
    p16:gp1 * EA2
    d16 pl2:f2
    DELTA1 cpd2:f3
    (p3 ph4):f2
    go=2 ph31
    d11 do:f3 mc #0 to 2
    F1EA(calgrad(EA1) & calgrad(EA2), caldel(d0, +in0) & calph(ph3, +180)
& calph(ph31, +180))
    4u BLKGRAD
    exit

ph1=0
ph2=0 0 2 2
ph3=0 2
ph4=0 0 0 0 2 2 2 2
ph31=0 2 0 2 2 0 2 0

;pl1 : f1 channel - power level for pulse (default)
;pl2 : f2 channel - power level for pulse (default)
;sp7: f2 channel - shaped pulse (180degree refocusing)
;spnam2: Crp140,1,20.1
;spnam3 : Crp60,0.5,20.1
;spnam7: Crp60comp.4
;p1 : f1 channel - 90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p3 : f2 channel - 90 degree high power pulse
;p16: homospoil/gradient pulse [1 msec]
;p24: f2 channel - 180 degree shaped pulse for refocusing
; = 2msec for Crp60comp.4
;d0 : incremented delay (2D) [3 usec]
```

;d1 : relaxation delay; 1-5 * T1
;d6 : delay for evolution of long range couplings (1 / 2J lr)
;d16: delay for homospoil/gradient recovery
;cnst13: = J(XH) long range
;inf1: 1/SW(X) = 2 * DW(X)
;in0: 1/(2 * SW(X)) = DW(X)
;nd0: 2
;ns: 2 * n
;ds: 16
;td1: number of experiments
;FnMODE: echo-antiecho
;use gradient files:
;gpnam1: SMSQ10.100

^{19}F - ^{13}C ^{n}J HMBC EXPERIMENT

```
;hmbcetgp_F19_f2.as
;
;This experiment is fluorine detected.
;

#include <Avance.incl>
#include <Delay.incl>
#include <Grad.incl>

"cnst30 = (1 - sfo2 / sfo1)/(1 + sfo2 / sfo1)"

define list<gradient> EA1 = ( 1.000 -cnst30 )
define list<gradient> EA2 = ( -cnst30 1.000 )

"p2 = p1 * 2"
"p4 = p3 * 2"
"d6 = 1s / (cnst13 * 4)"
"d11 = 30ms"
"d12 = 20u"
"in0 = inf1 / 2"
"acqt0 = 0"

"DELTA = p1 * 2 / 3.1416"
"DELTA3 = d6 - p14 - 4u"

# ifdef F1180
"d0 = 0.5 * in0
"DELTA1 = p12"
# else
"d0 = 3u"
"DELTA1 = p12 + d0 * 2 - 4u"
# endif

# ifdef CARBON
"DELTA2 = 2 * p16 + 2 * d16 + p24 + DELTA1"
"DELTA4 = ((d6 + DELTA3 + DELTA2) / 2) - d6 - p12 - 4u"
# else
```

```
"DELTA2 = 2 * p16 + 2 * d16 + p4 + DELTA1"
# endif

1 ze

# ifdef HDEC
d12 pl16:f3
2 d11 do:f3
# else
2 d11
# endif

    d12 BLKGRAD
    d1 pl1:f1

# ifdef CARBON
3 d12 cpds2:f3
    p1 ph1
    d6 pl8:f1 pl0:f2
    (p12:sp2 ph2):f1
    4u
    (p14:sp3 ph1):f2
    DELTA3 pl2:f2
    DELTA
    DELTA2 UNBLKGRAD

    (p3 ph3):f2
    d0
    (p12:sp2 ph2):f1 ;middle of t1 period
    d0
    4u do:f3
    p16:gp1*EA1
    d16 pl0:f2
    (p24:sp7 ph4):f2
    4u
    p16:gp1*EA2
    d16 pl2:f2
    DELTA1 cpds2:f3
    (p3 ph4):f2
```

```
# else
3 p1 ph1
    d6 pl0:f1 pl2:f2
    (center(p12:sp2 ph1):f1 (p4 ph1):f2 )
    d6
    DELTA
    DELTA2 UNBLKGRAD
    (p3 ph3):f2
    d0
    (p12:sp2 ph2):f1 ;middle of t1 period
    d0
    p16:gp1 * EA1
    d16 pl0:f2
    (p4 ph4):f2
    DELTA1
    p16:gp1 * EA2
    d16 pl2:f2
    4u BLKGRAD
    (p3 ph4):f2
# endif

# ifdef HDEC
    go = 2 ph31
    d11 do:f3 mc #0 to 2
    F1EA (calgrad(EA1) & calgrad(EA2), caldel(d0, +in0) & calph(ph3, +180)
& calph(ph31, +180))

# else
    go = 2 ph31
    d11 mc #0 to 2
    F1EA (calgrad(EA1) & calgrad(EA2), caldel(d0, +in0) & calph(ph3, +180)
& calph(ph31, +180))

# endif
4u BLKGRAD
exit

ph1 = 0
```

```
ph2 = 0 0 2 2
ph3 = 0 2
ph4 = 0 0 0 0 2 2 2 2
ph31 = 0 2 0 2 2 0 2 0

;pl1 : f1 channel - power level for pulse (default)
;pl2 : f2 channel - power level for pulse (default)
;sp7: f2 channel - shaped pulse (180 degree refocusing)
;spnam7: Crp60comp.4
;p1 : f1 channel - 90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p3 : f2 channel - 90 degree high power pulse
;p16: homospoil/gradient pulse [1 msec]
;p24: f2 channel - 180 degree shaped pulse for refocusing
; = 2msec for Crp60comp.4
;d0 : incremented delay (2D) [3 usec]
;d1 : relaxation delay; 1-5 * T1
;d6 : delay for evolution of long range couplings (1/2Jlr)
;d16: delay for homospoil/gradient recovery
;cnst13: = J(XH) long range
;inf1: 1/SW(X) = 2 * DW(X)
;in0: 1/(2 * SW(X)) = DW(X)
;nd0: 2
;ns: 2 * n
;ds: 16
;td1: number of experiments
;FnMODE: echo-antiecho
;for z-only gradients:
;gpz1: 80
;use gradient files:
;gpnam1: SMSQ10.100
```

^{1}H - ^{13}C - ^{19}F (3,2)D HCF EXPERIMENT

```
;HCF_2D.ry

#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>

"p2 = p1 * 2"
"p4 = p3 * 2"
"p22 = p21 * 2"
"d11 = 30m"
"d12 = 20u"
"d3= 1 / 2 * cnst10"
"DELTA1 = d3 - p14 / 2"                    ;cnst10 = J coupling 13C-1H
"d2 = (1 / 4 * cnst10)"
"DELTA2 = d2 - d16 - p16 - p14 / 2"
"DELTA4 = DELTA1 - p12"
"d4 = 1 / 4 * cnst11"                      ;cnst11 = J coupling 13C-19F
"d5 = d4 - d3"
"d0 = 4u"
"DELTA6 = d2 + d0 * 2 - p14 / 2"
"in0 = inf1 / 2"
"in33 = cnst5 * inf1 / 2"
"in43 = in33"
"TAU = p16 + d16 + 4u - p1 * 2 / 3.1416"

"d33 = d5 - (cnst17 * p24 / 2) - p12"
"d43 = d4 - cnst17 * p24 / 2"

"acqt0 = 0"
baseopt_echo

1 ze
2 d11
    d12 do:f3
    d12 pl3:f3
    d1 BLKGRAD
    d12 UNBLKGRAD
```

```
    (p21 ph5):f3                                ;1H 90 degree pulse channel 3
    d0
    (p21 ph1):f3                          ;1H 90 degree pulse channel 3 BIRDrx
    DELTA1 p10:f2 pl8:f1
    (center (p22 ph9):f3 (p14:sp3 ph1):f2 )                ;180 degree pulse
    (p12:sp2 ph1):f1
    DELTA4
    (p21 ph1):f3              1H 90 degree pulse on channel 3, end of BIRD
    d0
    DELTA2
    p16:gp1 * EA
    d16
    (center (p22 ph1):f3 (p14:sp3 ph1):f2 )
    DELTA6 pl2:f2
    (P21 ph2):f3                                          ;transfer to 13C
    (p3 ph1):f2
    d3 pl16:f3                          ;pl16 decoupling power channel 3
    d33 pl10:f2 cpd2:f3                          d33 is incrementing t2
    4u
    (p12:sp2 ph1):f1
    (p24:sp7 ph1):f2
    4u
    d43 pl2:f2

if "l0 %2 == 0"
    (p3 ph4):f2                                                    ;cos
    else
    (p3 ph7):f2                                                    ;sin

    d12 do:f3 pl1:f1
    p1 ph6
    TAU pl8:f1
    (p11:sp1 ph1):f1
    4u
    p16:gp2
    d16

    go = 2 ph31 cpd2:f3
    d11 do:f3 mc #0 to 2
```

```
    F1I(iu, 2)
    F1EA(calgrad(EA), caldel(d0, +in0) & caldel(d33, +in33) & caldel(d43,
-in43))
d2 BLKGRAD
exit

ph1 = 0
ph2 = 2
ph4 = 0 0 2 2
ph5 = 1 3
ph6 = 0 0 0 0 2 2 2 2
ph7 = 1 1 3 3
ph9 = 1
ph31 = 0 2 2 0 2 0 0 2

;pl1 : f1 channel - power level for pulse (default)
;pl2 : f2 channel - power level for pulse (default)
;pl3 : f3 channel - power level for pulse (default)
;pl0 : 0W decoupling power on channel 2
;pl8 : 0W decoupling power on channel 1
;pl16 decoupling power on channel 3 1H
;p1 : f1 channel - 90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p3 : f2 channel - 90 degree high power pulse
;p4 : f2 channel - 180 degree high power pulse
;p24 : 2ms BB refocusing 19F pulse
;p12 : 1ms BB inversion 19F pulse
;p14 : 500 us BB inversion 13C pulse
;p24 : 2ms BB refocusing 13C pulse
;spnam1 : Crp60comp.4
;spnam2 : Crp140,1,20.1
;spnam3 : Crp60,0.5,20.1
;spnam7 : Crp60comp.4
;d0 : incremented delay (2D)                                    [3 usec]
;d1 : relaxation delay; 1-5 * T1
;d11: delay for disk I/O                                        [30 msec]
;d12: delay for power switching                                 [20 usec]
;cnst5 = scaling factor
```

```
;cnst10 = 1JCH coupling 13C-1H
;cnst11 = J coupling 13C-19F
;cnst17 = -0.5 for Crp60comp.4
;inf1: 1/SW(H) = 2 * DW(H)
;in0: 1/(2 * SW(H)) = DW(H)
;nd0: 2
;ns: 2 * n
;ds: 16
;td1: number of experiments
;FnMODE: States-TPPI, TPPI, States or QSEQ

;use gradient ratio: gp 1 : gp 2
; 40 : 42.52 for F-19
```